

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 49.0254 Seconds
(without alignments)
4211.715 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMYYRR.....QHPKRVSQFLDDPSTAETVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5531.5	49.6	2201	2 A54774	ATP binding caset
2	2863.5	25.7	1529	2 A59189	ATP-binding caset
3	2694.5	24.2	1472	2 B54774	ATP binding caset
4	2503	22.5	1704	2 S71363	probable ATP-bindi
5	2502	22.5	1704	2 A59188	ATP-binding caset
6	1829	16.4	1802	2 T33783	hypothetical prote
7	1813	16.3	1816	2 A84845	probable ABC trans
8	1603	14.4	1317	2 C88925	protein F33E11.4 [
9	1600.5	14.4	1447	2 T15200	hypothetical prote
10	1408.5	12.6	1758	2 F88559	protein C48B4.4b [
11	1407	12.6	1767	2 S60124	transport protein
12	1405	12.6	1704	2 T42749	ATP-binding caset
13	1376	12.3	1246	2 T00826	hypothetical prote
14	1345	12.1	269	2 T46467	hypothetical prote
15	1038.5	9.3	1564	2 T27121	hypothetical prote
16	937.5	8.4	1431	2 T22748	hypothetical prote
17	901.5	8.1	373	2 T47150	hypothetical prote
18	741.5	6.7	1011	2 T07712	probable ABC-type
19	665	6.0	895	2 T07714	probable ABC-type
20	650	5.8	900	2 T07717	probable ABC-type
21	625.5	5.6	722	2 T07716	probable ABC-type
22	618.5	5.6	925	2 T07713	probable ABC-type
23	539	4.8	1336	2 T18288	ABC transport prot
24	435	3.9	582	2 H95950	probable ABC trans
25	429	3.8	330	2 S27707	daunorubicin resis
26	423.5	3.8	664	2 T07715	probable ABC-type
27	419.5	3.8	342	2 T36505	probable ABC-type
28	410	3.7	339	2 AE2102	ABC transporter AT
29	408.5	3.7	333	2 D72492	probable ABC trans

RESULT 1

AS4774
ATP binding cassette transporter ABC1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 02-Feb-2001
C:Accession: A54774
R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A:Reference number: A54774; MUID:94375008; PMID:9088782
A:Accession: A54774
A:Molecule type: mRNA
A:Residues: 1-2201 <LUC>
A:Cross-references: GB:X75926; NID:9495256; PIDN:CA53530.1; PID:9495257
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop
F:856-1047/Domain: ATP-binding cassette homology <ABC1>
F:873-880/Region: nucleotide-binding motif A (P-loop)
F:1869-2060/Domain: ATP-binding cassette homology <ABC2>
F:1886-1893/Region: nucleotide-binding motif A (P-loop)

ALIGNMENTS

30	408.5	3.7	339	2	S74048	probable daunorubi
31	403.5	3.6	326	2	JN0849	hypothetical 35K p
32	403	3.6	332	2	T36741	probable ABC-type
33	402	3.6	347	2	S76278	ABC-type transport
34	401.5	3.6	583	1	B64816	ABC-type transport
35	398.5	3.6	583	2	H90737	hypothetical prote
36	395.5	3.5	583	2	A85588	hypothetical prote
37	393.5	3.5	308	2	D97589	probable ATP-bindi
38	393.5	3.5	308	2	A12810	hypothetical prote
39	393	3.5	254	2	A11554	daunorubicin resis
40	393	3.5	325	2	S32908	hypothetical prote
41	392	3.5	196	2	T12512	hypothetical prote
42	390	3.5	314	2	D97318	ABC-type MDR trans
43	387.5	3.5	246	2	S75436	hypothetical prote
44	387.5	3.5	331	2	A52008	ATP-binding protei
45	386.5	3.5	257	2	C84147	ABC transporter (A

Query Match	49.6%	Score	5531.5	DB 2	Length	2201			
Best Local Similarity	49.4%	Pred. No.	0						
Matches	1107	Conservative	363	Mismatches	574	Indels	195	Gaps	27
QY	61	LPSAGTVPWLOGLICNVNNTCPQITPCGEPRLSNFNDLSVRSLLADARTVLGASAH	120						
DB	1	MPSAGTLEPWQGIICNANNPCFRYPTEGAPGVGNFNKSIIVSRILFSDAQRLLLYSQ	60						
QY	121	TLAGLGKLIATLRAAR	153						
DB	61	SIKMKHKLVRMLRQIKHPNSNLKQDLVDNETFSGFLQHNLSLPSRTVDSLLQXNVGLQ	120						
QY	154		160						
DB	121	KVFLQGYQLHLASLNGSKLEETIQLGDAEVSALCGLPRKKLDAARVLRYNMILKPVV	180						
QY	161	-----TSLRTESLGLAQOAPLHSLLEAABDLAQELLARLSVELR	212						
DB	181	TKLNSTSHLPTQHLA	233						
QY	213	TSGLPEL	265						
DB	234	SSSTQIIQVAVSRIVCGHPEGGLKIKSLNWDNNYKALFGGNTEEDVDTFYDNSTTP	293						
QY	266	ACSELIGALDGHPLSRLLWRRLKPLILGKLLFAPDTPFTRKLMQAVNRTFEEILLRDR	325						
DB	294	YCNDLMKNLESSPLSRRIIWKALKPLVGLKILYTPDTPATQVMAEVNKTQELAVFHDLE	353						
QY	326	EWEMLGPRITTFMNDSSNVAMLRQLLQW	377						
DB	354	GMWEELSPQIWTFMENSQEMDLVRLTLLDSRGNDQFWEQKLDGLDWTAQDINAFKPNPD	413						

C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

```
Query Match      25.7%; Score 2863.5; DB 2; Length 1529;
Best Local Similarity 41.1%; Pred. No. 3.7e-171;
Matches 644; Conservative 246; Mismatches 438; Indels 239; Gaps 36;

QY 745 GIATWYLEAVCGQVIGIPPMNPPRRSVCWGP-RPPKSPACPPTLPDPKVLV-----796
D 1 GILTWIEAVHPGMYGLPRWTFPLQKSYWLSGRTEAWSWPWARTPRLSVMEDQAC 60

QY 797 -----EAPPGCLSGVSVRSLEKRPFGSPQALRGLSLDFYQGHITAFGLHN 843
D 61 AMESRRFETRGMEETPLPLVVCVDKLTQVTKYDKDKLANKLSLNYENQVVSFLGN 120

QY 844 GAGKTTLSILSGLPSPGSGSFIIGHDVRSMAAIRPHLGVCPQVNVLFDMLTVDHVM 903
D 121 GAGKTTMSILAGFPPTSSGATIYGHDIRTEMDEIRKNGMCPQHNVLPDLRTVEEHL 180

QY 904 FYGRKLGLSAUVGPEDRLLDQVGLVSKQSVOTRHLGGMORKLSVAIAFVGGQVTL 963
D 181 FYSRLKMAQEIRREMDKWIEDLESNKRHSIVQTLGGMKRKLKLSVAIAFVGGRAIIL 240

QY 964 DEPTAGVDPASRGITWELLKYREGRTLLSTHLDDEALLGDRVAVVAGGRLCCGSP 1023
D 241 DEPTAGVDYARRAIWDLTKYKPGRTIILLTHMDEADLLGDRITAIISHGKLKCCGSP 300

QY 1024 FLRRHLGSGYILTVKARLPLTNEKADTDMBSVDTROEKNKGSGSRVGPQLLALVQ 1083
D 301 FLKTYGQGYRLTLVKR--PAEPGQEPQGLASSPPGRAPLSCE-----LQVQFIR 352

QY 1084 HWVPGARLVEPLHELVLVLPVYTGADHGSFATLPRDLRLAELRLTYGIGSDTSLERIF 1143
D 353 KHVASCLLVSDTSTELSYLPSEAAKGAFAERLFQHLERSLDALHLSPLGLMDTLEEVF 412

QY 1144 LKQVEE-----CAADT-----DMEDSGCGHLCGTGAGLDVTLRLKMPQPTALENPEP 1192
D 413 LKVSSEDSQLENSADVKESKDVLPFGAEGPASGEGHAG-NIARCELTQSQASLSQASS 471

QY 1193 AGSAPETDQSG-----PDVAG--RVQGMALTR-----1218
D 472 VGSA--RGDEGAGTYDYGDRPLFDNPQDPDQVNSLQVEAEALS RVGGSRKLDGWLKV 530

QY 1219 QQLQALLKRLFLARRSRGLPAQIVLPALFVGLALVSLIVPPGHPYALRLSPTMY--1276
D 531 RQFHGLLVKRFHCARRSKALPSQILLPAFFVCVAMTVALSVEIGDLPPLVLSPSQVHN 590

QY 1277 -----CAQVSFSED-----APQDPCRALLLEALLQEAG-----LEEP-----PV 1311
D 591 YTPQGRNFIYANEERREYRLRLSP-DAS PQQLVSTFRLPSGVGATCVLKSPANGSLGPT 649

QY 1312 QHSSH-----RFSAPVPAEVAKVLASGNWTPESPSPA-----1344
D 650 LNLSGESRLLAARFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPASDEDLQANVSL 709

QY 1345 -----CQCSQPGARRLLPDCPAAAGGPPPPQAVTGSGEYVQ 1380
D 710 PPTAGPEMWTSAPLRLVREVRCTCSAQGTGF---SCPSSVGGHPQMRVV-TGDILT 765

QY 1381 NLTGHNLSDFLVKTYPRLVROQLKTKKWNVEYCGPSLGGDRDPLGSGQELGRSVLEWL 1440
D 766 DITGHNVSEYLLFTSDRFLH-----RYGAITFG-----794

QY 1441 ALLSPLPGGALDRVLKNTA-----WAHSLDAQDSIKTFNNKGMHSMVAFVNRASN 1492
D 795 -----NVUKSIIPASFGTRAPWVRKIAVRAAQVFYNNKGHSMPTYLSLNN 842

QY 1493 AILRAHLPP---GPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSLVCVWFAMSFV 1549
D 843 AILRLANLPKSKGNPA--AVGITVTHNPMNKTSSASLSLDYLL-QGTDVVIATFIIVAMSFV 899

QY 1550 PASFTVLVIBERVTRAKHLQMLGGSLPITYLWGNFLWDMCNVLPACIVLFIPLAQORA 1609
D 900 PASFVVFLVAEKSTAKHQLQFVSGCNPIIYWLANTYVMDMLNLYLPATCCVILFVFDLPA 959
```

```
QY 1610 YVAPANLPALLLLLYGMSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVL 1669
D 960 YTSPTNFPFVLSLFLYLGMSITPIMYPASFWEVPSAYVFLVINLFIGITATVATFLL 1019

QY 1670 ELFS--DQKLEQVSRILKQVFLIPHPFCILGRGLIDMVNRQAMADAFERLGD-RQFQSPLRW 1727
D 1020 QLFEHDKDLKVNSVLSKSCFLIFPNYNLGHGLMAYNEYINEYAKIGQFQDKMKPFWE 1079

QY 1728 EVVGKNLAMVLOGPLFLFTLLLOH-----RQLLPQPRVRSLSPLLGEDEDEVARERVR 1783
D 1080 DIVTRGLVAMAVEGVVGFLITIMCQYNFLRRPQMP---VSTKPV--EDVDVVASRQVR 1134

QY 1784 VQAGTQGVVLRLNLTQVYRGO---RMPAVDRCLCLGIPPGCEGFLGVNGAGKTSFRMV 1840
D 1135 LRGDADNDMVKLENLTQVYKSRKIGRILLAVDRCLGVRFGECEGFLGVNGAGKTSFTKML 1194

QY 1841 TGDITLASGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGTREHLLELRLGRVPEA 1900
D 1195 TGDESTTGGCAFVNGHSLKELLQVQSLGYCPQCDALPDELTAAREHLQLYTRLRGISWK 1254

QY 1901 QVAQTAGSLARLGLSWYADRPAGTYSNGNKEKLATALVGDPAVFLDEPTTGMDPSA 1960
D 1255 DEARVVRKWALEKLELTQYADRPAGTYSNGNKKLSTATAALIGYPAPFIDLEPTTGMDPKA 1314

QY 1961 RRFLWNSLLAVVREGSRVMTLSHSMEECEALCSRLAIVMNGRFRCLGSPQHLKGRFAAGH 2020
D 1315 RRFLWNLILDLIKTGRSVVLTSHSMEECEALCTRLAIVMNGRLRCLGSLQHLKNRPGDY 1374

QY 2021 TLTLRVPAARS-QPAAAFVAAEFPGSELREAHGRLRQLPFGGRCALARVGEIAVHGA 2079
D 1375 MITVTKSSQSVKDVVRFNRRNFPPEMLKERHHTKVQVQL-KSEHISLAQVFSKMEQVSG 1433

QY 2080 EHGVEDPSVQTMLEEVFLYPSKDGKDEDETEQKEAGVGVDPAQGLQHPKRVVSQPLDDP 2139
D 1434 VLGIEDYSVQTLDNVFNFAK---KQSDNLEQOE---TEPPSALOSPLGCLLSLRP 1486

QY 2140 STAETVL 2146
D 1487 RSAPTEL 1493

RESULT 3
BS4774
ATP binding cassette transporter ABC2 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 16-Aug-2004
C:Accession: BS4774
R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chimini, G.
Genomics 21, 150-159, 1994
A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
A:Reference number: AS4774; MUID:94375008; PMID:8088782
A:Accession: BS4774
A:Molecule type: mRNA
A:Residues: 1-1472 <LUC>
A:Cross-references: GB:X75927; NID:g495258; PIDN:CAA53531.1; PID:g495259
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:44-234/Domain: ATP-binding cassette homology <ABC1>
F:61-68/Region: nucleotide-binding motif A (P-loop)
F:1108-1300/Domain: ATP-binding cassette homology <ABC2>
F:1126-1133/Region: nucleotide-binding motif A (P-loop)

Query Match      24.2%; Score 2694.5; DB 2; Length 1472;
Best Local Similarity 40.8%; Pred. No. 1.4e-160;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

QY 796 VEEAPPGLSPGVSVRSLEKRPFGSPQALRGLSLDFYQGHITAFGLHNGAGKTTLSILS 855
D 16 MEEEPHTLPLVVCVDKLTQVTKYNDKKLANKLSLNYENQVVSFLGRHNGAGKTTTWSIUT 75

QY 856 GLPPSPGGSFALGHDRVSSMAAIRPHLGVCPQVNVLFDMLTVDHVMFYGRKLGSAAV 915
D 16 MEEEPHTLPLVVCVDKLTQVTKYNDKKLANKLSLNYENQVVSFLGRHNGAGKTTTWSIUT 75
```

Db 76 GLPPTSGSATYCHDIRTEMDEIRKXNLGCMCPQHNVLFDRLTVEHLLWFYSRKLSMAQEE 135
Qy 916 VGPQDRLLQDVGLVSKOSVQTRHLSGMQRKLSVAIAFVGGSQVWILDEFTAGVDPASR 975
Db 136 IRKETDWMIEDLELSNKRHSLVQTLSGGMRKLSVAIAFVGGSRAILDEFTAGVDYAR 195
Qy 976 RGIWELLKYPREGTLLSTHLDAAELGDRVAVAGGRLCCGSPFLFRRHLGSGYYL 1035
Db 196 RAIWDLILKYKPGRTLLSTHMDAADLLDRIAILSHGLKCCGSPFLKGYXGYRL 255
Qy 1036 TLVKARPLTNEKADTMDEGSDVTRQCKNGSGSRVGTPOQLLALVQHWVPGARLVEEL 1095
Db 256 TLVKQPAEPGTSQBPGLASSPGCPRL-----SSCSBPQVSQFTRKHVASSLLVSDT 307
Qy 1096 PHELVLVLPYTGADHSGFATLFRDLTRLAELRLTYGIGSDTSLEELFLKVVE----- 1149
Db 308 STELSYLILPEAVKGAFLFQLEHSLDALHLSFGLMDTTLVEEVLKVSSEDSLEN 367
Qy 1150 CAADT-----DMEDSGCGHLCGTGIAGLDVTLRLKMPPEQTALENGEPAGSAPETDQSG 1204
Db 368 SEADVKSERKDVLPFAEGLTAVGGQAG--NLARCSELAQSQASQSASVSGA--RGEETG 425
Qy 1205 -----PDVAG-----RVQGWALTRQOLQALLKREFL 1230
Db 426 YSDGYGYRPLFDNLQPDNVLSLOEAMEALAQVCGQSRKLEGGWLLQKRGHGLLVKRFH 485
Qy 1231 LARRSRGLFAQIVLPALFVGLVGLVFLVPPFGHYPALRLSPMY-----GAQVSFF 1283
Db 486 CARENKALCSQILLPAFFVCVMTVALSVPEIGDLPLVLSPSQYHNYTPQRCGNFPIYA 545
Qy 1284 SED-----APQDQGRALLEALQEG-----LEEP-----PVQHSSH----- 1316
Db 546 NEERQERYRLRSP--DASPQQLVSTFRFLPSGVGATCVLKSPANGSLGPMNLNSGESLLA 604
Qy 1317 -RESAPEVPAEVAKVLASGNWTPESPSPA----- 1344
Db 605 ARFPDSCLESFTQGLPLNSVPPPPPPAPSDSPVXPDEDSLQWNMSLPTAGPETWTS 664
Qy 1345 -----CQCSQPGARRLLPDCPAAGGPPPPQAVGTSGEVVQNLTRGNLSDFL 1391
Db 665 APSLPLRVHVPVCTCSAQGTGF--SCPSSVGGHPQMRV--TGDILTDTGHNVSEYL 720
Qy 1392 VKTYPRLVRQGLTKKWNVRYGFSLGGDRDGLPS--GQELGRSVEELWALLSLPLPGG 1449
Db 721 LFTSDRFRHL-----RYGAITFGNVKQSPAPASFGARVPPMVRKI----- 759
Qy 1450 ALDRVLKNTAWAHSLDAQDSLKTFWKNKGWMSVAFVNRASNAILRAHLPP---GPARH 1506
Db 760 AVRRV-----AQVLYNKGVHSMPTVYLSNNAILRANLPKSKGNPA-- 801
Qy 1507 AHSITTLNHPNLTRKEQLSEALMASSVDVLVSVICVVPFAMSFPVSPASFTLVLIBERVTRAK 1566
Db 802 AYXITVTNHPNKTSSASLDLYLL--QGTVDVIAIFIIVAMSGFVPASFWVFLVAEKSTKAK 860
Qy 1567 HLQMGGLSPPLYWGLNPLWDMCMYLPACIVLVIFLAFQORAVVAPANLPALLLLLLLY 1626
Db 861 HLQFVSGCNPIYWLANYVMDMLNLYLPATCCVILFVFDLPAYTSPNTPFAVLISFLY 920
Qy 1627 GWSITPLMYPASFFSPSTAYVVLTCINLFIGINGMATVFLVLFSS--DOKLQVSRILK 1685
Db 921 GWSITPLMYPASFFSPSTAYVFLVINLFIGITATVATFLLQLQFHEHDKDLKVNSYLK 980
Qy 1686 QVFLIFPHFCILGRGLIDMVRNQAMADAFERLGD--RFQSPULRWEVVGKNLLAMVIOGFLF 1744
Db 981 SCFLIFPNYNLGHGLMEMAYNEYINEYAKIGQFDKMSPPPEWDIVTRGLVAMTVEGFVG 1040
Qy 1745 LFLFLLLOHRSQILLPOPR---VRSPLPLGDEDEDAVERERVRVCGATQGVVLVRLNTKV 1801
Db 1041 FFLTIMCOY--NFLRQFORLVPSTKPV--EDVDVVASERQVRLRGDADNDNMVKIENLTKV 1096
Qy 1802 YRGO---RMPAVDRILCLGI--PFGSCFGLLVNGAGKTSITPRMVTGDTLARSGEAVLAGHS 1857
Db 1097 YKSRKIGRILAVDRILCLGVCVPGECFGLLVNGAGKTSITFKMLTGDSSTTGGEAFVNGHS 1156

Qy 1858 VAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSW 1917
Db 1157 VLKDLQVQOSLGYCPQFDPVDELTAAREHLQLYTRLCIPKWDQEAQVVKWALEKLELTK 1216
Qy 1918 YADRPACTYSGGNKRKLATALALVGDPAVVFLEPTTGMDDPSARRFLWNSILLAVVREGRS 1977
Db 1217 YADKPACTYSGGNKRKLSTALIALIGYPAFIFLEPTTGMDDPKARRFLWNLILDILKTGRS 1276
Qy 1978 VMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARS--QPAAA 2036
Db 1277 VMLTSHSMEECEALCRLAIMVNGRLHCLGSIQHLKNRFGDGYMITVTRKSSQNVKDVVR 1336
Qy 2037 FVAAPFGSELRAHGG-----RLRFQLPFGRCALARVFGELAVHGAHGVEDFSYSQT 2091
Db 1337 FFRNRP-----BAHAQGGKTPYVQVQL--KSEHISLAQVFSKMEQVVGVGLGIEDYSVSQT 1390
Qy 2092 MLEEVFLYFKDQCKDETEEQEAGVGVDPAP-----GLQHPKRVSQFL-----DDPSTA 2142
Db 1391 TLONVFNFAK---KQSDNVEQEAEPSSLPGLGLLSLLRPPAPTELRALVADEPEDL 1447
Qy 2143 ET 2144
Db 1448 DT 1449

RESULT 4
S71363
probable ATP-binding cassette transporter ABC-3 - human
N/Alternate names: ATP-binding cassette transporter ABC-3
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 16-Aug-2004
C/Accession: S71363
R/Klugbauer, N.; Hofmann, F.
FEBS Lett. 391, 61-65, 1996
A/Title: Primary structure of a novel ABC transporter with a chromosomal localization on
A/Accession: S71363
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1704 <KLU>
A/Cross-references: UNIPROT:Q95758; EMBL:X97187; NID:g1514529; PIDN:CAA65825.1; PID:e243
A/Experimental source: cell line medullary thyroid carcinoma
C/Genetics:
A/Gene: GDB:ABC3
A/Cross-references: GDB:3770735; OMIM:601615
A/Map position: 16p13.3-16p13.3
C/Superfamily: ATP-binding cassette binding: P-loop; phosphoprotein; transmembrane prote
C/Keywords: ATP binding; nucleotide binding; P-loop; phosphoprotein; transmembrane prote
F/255-283/Domain: transmembrane #status predicted <TM1>
F/307-329/Domain: transmembrane #status predicted <TM2>
F/345-364/Domain: transmembrane #status predicted <TM3>
F/373-394/Domain: transmembrane #status predicted <TM4>
F/401-422/Domain: transmembrane #status predicted <TM5>
F/452-475/Domain: transmembrane #status predicted <TM6>
F/549-739/Domain: ATP-binding cassette homology <ABC1>
F/566-573/Region: nucleotide-binding motif A (P-loop)
F/685-690/Region: nucleotide-binding motif B
F/1100-1120/Domain: transmembrane #status predicted <TM7>
F/1145-1169/Domain: transmembrane #status predicted <TM8>
F/1181-1207/Domain: transmembrane #status predicted <TM9>
F/1215-1236/Domain: transmembrane #status predicted <TM10>
F/1245-1264/Domain: transmembrane #status predicted <TM11>
F/1299-1324/Domain: transmembrane #status predicted <TM12>
F/1399-1590/Domain: ATP-binding cassette homology <ABC2>
F/1416-1423/Region: nucleotide-binding motif A (P-loop)
F/1535-1540/Region: nucleotide-binding motif B
F/674,866,1524/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #stat
F/1344/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predi

Query Match 22.5%; Score 2503; DB 2; Length 1704;
Best Local Similarity 33.7%; Pred. No. 1.9e-148;
Matches 639; Conservative 313; Mismatches 631; Indels 314; Gaps 47;

QY 307 LMAQVNRTEELTLARDVREVMELGPRIFTFNMDSNNVAMLQRLQMQDEGRQPRPGG 366
D 9 LLLKXNYTLQKQKVLVTLEFL---PLLF-----SGILWLRLKIQSENVENATYFG- 59
QY 367 RDHMEALSLFLDPSGGYSWQ---DAHADVHLVGTQLGRVTECLSLDKLEAPRAALV 422
D 60 -OSIQELPLFFTPFPDGTWEALYPSHSDAAKTVTETVRRALVINM-RVRGFPPEKDFE 117
QY 423 SRALQLLAEBHRFWAGVFLGPDSDPTBHTPDLGPGHVRKIRMDIDVTR----- 475
D 118 DVIRDNCSVLAADVPEHPNHS---KEPLPLAVKYLHRSYTRNTYMTQTGSFFLK 174
QY 476 -----TNKIRDFWDPGPAADPLT-----DLRYVMGGFYVLQDLVERAAVRLSGAN--- 522
D 175 ETEGWHITSLFLFPNPGP-REPTSPDGEGPCYIREGFLAVQHAVDRAIMEYHADAATRQ 233
QY 523 --PRAGLYLOQMPYCYVDVFLVLSRLPLFLFLMIYSVTLTVKAVRKEKTRLDRT 580
D 234 LQRLTVTIKRFYPPFTADPFLVAIQQLPLLLLLSFTYTALTITARAVVQEKERLREY 293
QY 581 MRAMGLSRAVLWGLNFLSLCLGPFLLSALLVLVLKLG-----DILPYSHPGVVFLFLAAF 635
D 294 MRAMGLSSLWHSANFLFLFLLIAASPMTLFCVKVKNVAVLSRSPSLVLAFLCF 353
QY 636 AVATVTSQFLLSAFTSRANLAAACGLAYFSLYPLVLCVAVWRDRLPAGGRVAASLLSPV 695
D 354 AISTISFSFVSTFTSKANMAAARGFLYPTFYIPFVAPRYNMTLSQKLSCLSNV 413
QY 696 ARGFCESLALLEBEGQAHNVGTRPTA-DVFSLAQVSGLLLDAAALYGLATWYLRV 754
D 414 ANAMGAQIIGKPEAKGMGIQMRDLSLVNVDVDFCGQVLGMLLDVLYGLVTVMEAV 473
QY 755 CPQYGIPEPNPPRRSVMCCPRPKSPACPTPLDPKVLV-----BEAPGLSPGVSV 809
D 474 PPGQGVQVQWTFFTMPSWCQ-KPRVAGKEEDSDPEKALURNEYFAEPEDLVAGIKI 532
QY 810 RSLEK--RFPQSPQALRGLSLDFYQHHTAFGLHNGAGKTTLSLGLFPPSPGSAFI 867
D 533 KHLKSVFRVGNKDRAVRDLNLNLYEGQITVLVGHNGAGKTTLSMLTGLFPPTSGRAYI 592
QY 868 LGHDVRSMAAIRPHLGVCPQNVNLFDMLTVDHWFVGRKLGKLSAAVVGPDRLQDV 927
D 593 SGYESQDMVQIRKSLGLCPQHDILFDNLVTAELHYFYAQLKGLSRQCPPEVKQMLHII 652
QY 928 GLVSKQSVQTRHLSGMORKLSVATFVGGSOVTLDEPTAGVDPAARRGITWELLKLYRE 987
D 653 GLEDKNSRSLSGMRKLSIGALTAGSKVLILDEPTSGMDAISRAIWDLLQOKS 712
QY 988 GRTLSTLTHLDEALLGDRVAVVAGRLCCCGSPFLFLRRHLGSGYTLTVKARLPLATN 1047
D 713 DRTVLTTHFMDADLLGDRIAMAKGELQCCGSSFLKQKYGAGYHMLVKE----- 765
QY 1048 EKADTMEGSDVTRQKNGSGSVGTPQLLALVQHVWPGARLVEELPHELVLVLPYTG 1107
D 766 -----PHCNPEDISQLVHHVHPNATLSSAGAEISFILPRES 802
QY 1108 AHGSGPATLFRDLTLAELRLTGVIQDTSLEELFLKV----- 1146
D 803 TH--RFEGLPAKLEKQKELGASFGASITTWEEVFLRVGKLVDSMDIOAIQLPALQVQ 860
QY 1147 VEECAADTDMEDGSCQHLCTGIAGLDVTLRLKMPPEQTALENGBPAGSAPETDQSGPD 1206
D 861 HERRASDWAVDNLCG-----AMDPSDGI GALIEER-----T 893
QY 1207 AVRGVQWALTRQOQALLLKRLFLARRSRGLFAQIVLPALFVGLALVSLVPPFGHY 1266
D 894 AVKLTGLALHCQQFWAMFLKKAAYSREWKMVAQVLVPLTCVTLALLAINYSSELFPD 953
QY 1267 PALRLSPMYQVQVFFSDEPDGCRARLLEALLQEAAGLEPPVQVHSHRFSAPVPAE 1326
D 954 PMLRLTGEYGRVTVFVSF-----PQTSQQLQOOL-----SEH 985

QY 1327 VAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPPPQAVTSGEVVQNLTGRN 1386
D 986 -----LKDALQAEQ--QSPREVLG-----D 1003
QY 1387 LSDFLVKTYPRLVROGLKTKKVVNEVRYGGFSLGGRDPLGSPGOELGRSVEELWALLSPL 1446
D 1004 LEEFLI-----FRASVEGGGF----- 1020
QY 1447 PGAGDLRVLKNLTAWAHSIDAQDSLKIWFNNKGMHSMVAFVNRASNAIIRAHLPDGPARG 1506
D 1021 -----ERCL--VAASFRTDVGERTVYNALFNNOYHSPATALAVDNLKFK--LLCGP--H 1069
QY 1507 AHSITTLNHP-----LNLTKQESLAALMASSVDVLVSVVVFAMSVFPASTFLVLIBER 1561
D 1070 A-SIVSVNFQPSRALQAQKQFNEG---RKGFIDALNL--LFAMAFLASTESILAVSER 1123
QY 1562 VTRAKHLQMLGSLSTLYLGNFLWDMCNLYLPACIVLVILFLAFOORAVVAPANLPALL 1621
D 1124 AVQAKHVQFVSGVHVASFWSALLWDLISFLPSLLLVVFKAFDVRAFTRDGHMADTLL 1183
QY 1622 LLLLYGWSITPLMYPASPFPSVPSTAYVVLTCINLFIGINGSMATPVLEL---FSDQKQ 1678
D 1184 LLLLYGMAIIPMYLWNPFFLGAATAYTRLTIFNLISGI---ATFLMTWIRIPAVKLE 1239
QY 1679 EYSRLKQVFLIFPHCLGRGLIDMVNR-----QAMADAFERLGRDQFQSP-LRWE 1728
D 1240 ELSKTLDHVFLVLPNHCLGMAVSSPVENVETRRYCTSSSEVAHYCKYNIQIENFYAS 1299
QY 1729 V--VGNLLAMVIQ-----PLFLFTLLQHSOLLPOQVR-----SLPLIGED 1773
D 1300 APGVGRFVASMAASCAYLILLFLTETNLLQRLGILCALRRRTLTLYTRMPVL-PED 1358
QY 1774 EDVARERERVQATQG---DVLVLRLNTKVYRGORMP---AVDRLCIGTPGCECFLGLV 1828
D 1359 QVADERTRIAPSDPSLLHTPLIILKLSKVTE-QRVLLAVDRLSLAVKQCEGFLG 1417
QY 1829 NGAKTSTFRMVGTTLASGEAVLAGHSVAREPSAAHLSMGYCSQSDAIFELLTGREHL 1888
D 1418 NGAKTTFKMLTGESLTSGDAFVGHRISSDVKVQRQICGCFQFDALLDHTMGREWL 1477
QY 1889 ELLARLGRVPEAQVATAGSLARGLSWYADRPACTYSGGNKRKLATLALVGDPAVVP 1948
D 1478 VMYARLGRIPERHIGACVENTLRGLLEPHANKLVRTYSGGNKRKLSTGIALIGSPAVIF 1537
QY 1949 LDEPTTGMDSARRFLWNSLLAVVREGSVMLTSHSMEECEALCSRLAIVNNGRFRCLGS 2008
D 1538 LDEPSTGMDPVARRLLMDTVARARESGNAIITSHSMEECEALCTRLAIVMOGQFKCLGS 1597
QY 2009 PQLKGRFAAGHTLTLRVPAARSQPA-----AAFVAAEFPGSSELRAHGRLRFLQPPGGR 2064
D 1598 PQLKSKFGSGVSLRAKVSQEQEALBEFAFVDLTFPGSVLEDEHQGMVHYHL-PGRD 1656
QY 2065 CALARVFCGLAVHGAHGVDFSVQTMLEEVFLYFS 2101
D 1657 LSWAKVFGILKAKEKYGVDDYSVQISLEQVFLSFA 1693

RESULT 5

A59188

ATP-binding cassette transporter ABC3 - human

C:Species: Homo sapiens (man)

C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004

C:Accession: A59188

R:Connors, T.D.; van Raay, T.J.; Petry, L.R.; Klinger, K.W.; Landes, G.M.; Burn, T.C.

Genomics 39, 231-234, 1997

A:Title: The cloning of a human ABC gene (ABC3) mapping to chromosome 16p13.3.

A:Reference number: A59188; MUID:97179225; PMID:9027511

A:Accession: A59188

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1704 <CON>

A:Cross-references: UNIPROT:Q99758; GB:U78735; NID:g1699037; PIDN:AAC50967.1; PID:g1699037

C:Genetics:

A;Description: The sequence of *C. elegans* cosmid Y39D8C.
A;Reference number: Z21408
A;Accession: T33783
A;Status: Preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1802 <BEC>
A;Cross-references: UNIPROT:Q9TXV8; EMBL:AF101313; PIDN:AAC69223.1; GSPDB:GN000023; CESP:
A;Experimental source: strain Bristol N2; clone Y39D8C
C;Genetics:
A;Gene: CESP:Y39D8C.1
A;Map position: 5
A;Introns: 45/3; 114/1; 195/1; 230/3; 543/3; 794/1; 849/1; 1036/2; 1099/1; 1132/3; 1165/
C;Superfamily: ATP-binding cassette homology

Query Match	16.4%	Score 1829;	DB 2;	Length 1802;
Best Local Similarity	28.4%;	Pred. No. 4.2e-106;		
Matches	539;	Conservative 312;	Mismatches 667;	Indels 382; Gaps 55;
QY	371	EA LRSFLDPGSGYSGWDAHADYCHLVGLTGRVTECLSLD---	KLEAAPSEAA LVS---	R 424
Db	124	DATKQIMDKTQNRYTANDLLNPVAAAYTKGLYNTAVFLNTNMTYKGTTEGEVMSVMQO 183		
QY	425	ALQLLAEHRFWAGVYVFLGPDSDSPTHEPTPDLGPGHVRIKIRMDIDWVTRTN---	KIR 480	
Db	184	QFQSECDNPLLAGIVF---DDSIADKLKNDP-----KRDFTYITRLNTHRRSR 229		
QY	481	DRFWDPGPAADPLTD--LRYVWG-----GFVYLQDLVERAAVRVLSGA 521		
Db	230	NAFGDSYPWDTSVPAQVYVSGP INPDDNDGSGPGVQWQEGFMTVQRAVDVAITEITGCE 289		
QY	522	N----PRAGLY-IQOMPYPICYDDVFLURLSRSLPLFLTLAWIYSVTLVKAVVRKEKTR 576		
Db	290	DAQLTPLLDSDYSQSRFPFGYSTKI-IEIGAFPMPIVIFSFMTSYIYIVRAVVVVEKDR 348		
QY	577	LROTBRAMGLSRAVLMLGFWFLSCLGPELLSAAILLVVLKGLDILPYSHPGVWVFLAAFA 636		
Db	349	LKEYRMVGLSQFINWVAHI--INTAKLTFVAVITILMHFVALKSDMTLMEVFVLMIIYA 406		
QY	637	VATVTOSFLSASFPSRANLAACGGLAYFSLYLPLYVL--CVAMRDRLPAGGRVAAASLLSP 694		
Db	407	FDVVYPAFMISSPMNSATSLISVVFWMLLYFWYAFSSIDQTNYPPLGYRLINCINPD 466		
QY	695	VAFPGCESLALLEBQEGAGQWNVGTRPAD---VPSLAQVSGLLLLDAAVLGLATWYL 751		
Db	467	IALNYGLQLAAAYETOADGLKWELETPPSPDNMLTFGHALIA--LIVDGIIMILTWYI 524		
QY	752	EA VCPGQYGIPE-PMNFPFRSVMCGPRPKSPACETPLD-----PKVLVEAPPG 802		
Db	525	EA VIPGEGVPQKPFVFL-PSYWF---PNSGSKTVDSDDQFQI QYADHVKLEKEPTD 579		
QY	803	LSPGSVRSLEKRFPGS-----POPALRGLSLDPYQCGHITAFILGHNAGKGT 848		
Db	580	LPI TINVNLTKTYGTSFPFKKLFDCFKGSGKEKRAVSNLNLKMYPGQCTVLGHNAGKGS 639		
QY	849	TTLSILSGILFPFGGSAFILGHDVRSMAAIRPHLGVCPQYNVLFDMLTVDHEVWFYGLR 908		
Db	640	TTFSMLTGVASPPSGSAYVNDFDIRTSLPKIRREMGLCPQYNTLFGFMTVMEHLBFFAKL 699		
QY	909	KGLSAAVVGPEQDR-LLODVGLVSKOSVQTRHLSGGMQRKLSVAIAFVGQSQVWILDEPT 967		
Db	700	K---ERTWDPPEAREILARIDFKADFAGALSGGQKKKLSLAIALIGSEVYVMDLEPT 756		
QY	968	AGVDPASRGIWELLKLYREGRTLILSTHLDABELLGDRVAVVAGGRLLCCGSPFLRR 1027		
Db	757	SGMDPGARHETWLIQREKERRITLTHFMEADLLGDRIAMAGQLCECCSPWFLKQ 816		
QY	1028	HLGSGYYLTLVKARLPLTTNEKADTDMGSDVTRQEKNGSGSRVGTGPQLLALVQHWPV 1087		
Db	817	QYGDGYHLTIV-----YDTSSTPDVSKTTD-----IIREYIP 848		
QY	1088	GAALVEELPHELVLVLPYTGADGGSFATLPRELDTRLAELRLTGYISDTSLEEIFLKV 1147		
Db	849	EAHFVSYIGOEATYLLSAT--HRPIPKFLKLELDEHQTCCGTSFGVSTITMEEVFLKVG 906		

QY	1148	BECAADTDMEDSGCQHCLCTGIAGLDVTILRKMPQPOTALENGEPAGSAPETDQGGSPD-	1206
Db	907	HTADERINYEHG-----IEN---DISEMIEKDDPILQDL	937
QY	1207	-AVRGVQWALTRQOQALLLKRELLARRSERGLFAQIVLPALFVGLALVPSLIVPPFGH	1265
Db	938	RAQVRVTGFTLOMOHAKAMFYKRAIFFRKWTQFLPOLVFPVAYLYLVMVFTSQVLP----	993
QY	1266	YPALLRSPTMYGAQVSFSEDAPGDCGRARLLLEALQEAGLEBPVQVHSHRFSAPRVA	1325
Db	994	-----SVKE-----	997
QY	1326	EVAKVLASGNWTPESPSPACOCQOPGARRLLPDCPAAAGPPPPQAVTSGSEVGVNLTG-	1384
Db	998	-----QDPQTISLAPFSDTKAG-----HLVSDSGNVVTLGGG	1031
QY	1385	RNLSDFLVKTYPRL-VROGL-----KTKKWNVREYGGFSLGGRDPLPGQBLGRSVEE	1438
Db	1032	QNLSSVMQGVTTQLGVTTVVDTISNVEKFIQDNTN---AMGSRITFGLHYA-----	1079
QY	1439	LWALLSPLPGGALDRVLKNTLTAWAHSLDAOQSLKIMFNNGKMSWAFVNRASNAILRAH	1498
Db	1080	-----LGFPV-----SMFNFSVSVPSLKTFFNNFGLYTPALAITFTDSMLSQ-	1123
QY	1499	LPPGPARHAHSITTLNHL-----NLTKQLSEAA--LMASSVDVLVSICVVFAMSFVP	1550
Db	1124	-----KOKQOYSFTAVNHLPPSTQDTLONTNRSDGAAFLIAYGLIVSFACV-----	1171
QY	1551	ASFVLVLIERVTRAKHLQMLGGLSPTLYWLGHFLWDMCNLYVPACIVVLIFLAFQORAY	1610
Db	1172	AGYSQFLITERKCKSKHQLSGIRPWFMEFLTAFIWDAAWFMVIRILCFDAIFYFINITAY	1231
QY	1611	VAPANLPALLLL-LLYLGWSITPLWYPASPFERSVPTAYVVLTCINLFTINGSMATFVL	1669
Db	1232	THDPGVMLIUTLSELLYLGWTALPFTWFQPFESAPKGMFMVWYHILTMIGISIAVPIL	1291
QY	1670	E-----LFSDKLOEVSRI-----LKQVFLIPFH-----CLG	1697
Db	1292	SQTSLLDAGYLSIIFAWLPPTNYISQIATVTFQENVRILACKKLDCTIPMKRAVTACCG	1351
QY	1698	RLIDWNRQAMADAFERL-----GRQFOSPLRWEVVGKMLAMVI-----QQLPL	1745
Db	1352	-----TASERLYVDNVLFVGNR-----KGILVYIFLAVQGFYIW	1386
QY	1746	LFTLLLOHR--SOLLQOPVRSL--PLLGEEDVARERE-----RVQOGATQGD	1791
Db	1387	IWFMRENDOFTKLPAIRCKADNPIDWITDTDKVDERDESDVIAEKSVVQRANNN	1446
QY	1792	--VLVLNLTKVYRGORMPAVDRLCLGIPGECFGLLVNGAGKTSFRMTVGTDLTASRG	1849
Db	1447	KTALVSNLIVKY--GNFNAVGVNPHVNSKDCFGLLVNGAGKTSFTQWLGTGENSISSG	1504
QY	1950	EAVLAGHSVAREPSAHLNGYCPQSDAIFELLTGREHLELILARLGRVPEAQVAQTAGSG	1909
Db	1505	DAYVNGSVKNNWREAGANTGCPQYDAIIKEMSGSETLYMFARIRIGPEKDIPIKKVNAV	1564
QY	1910	LARILGSWADR PACYTSNGNKEKLATALACDPVAVFLDEPTTCMDPSARRFLWNSLL	1969
Db	1565	IHAIGIMYASROIKTYSGNKKRLSIGIAVLGPLVLLDEPTSGVDPKARRIIWNILN	1624
QY	1970	AVVREGRSVMLTSHSMBEECALSRILMVNGFRCLGPSQHLKGRFAAGHTLTLRVPA	2029
Db	1625	RLRDLGTALVLTSHSMBEECALCTELAIWYKFCYGCQHIKSYGSGYTLILFLKNR	1684
QY	2030	R-SOPAAAFVAABFPGESELRHAGBLRFQLPPGRCALARYVFGELAVHGAHBGVDFSV	2088
Db	1685	NDAEKTSTIKQTFRGSVIKEEHLQNLNIDPRGD-SWSRLFKEKLETYSTSLNDDYSL	1743
QY	2089	SQTMLEVEFLYPSKDOGBDETEEQEAGVGVDPAFGLQH	2128
Db	1744	SQTTLSEQVLEFREDRAGVSSDSFSDGASSTGSGANSRQON	1783

```
RESULT 7
A84845
probable ABC transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84845
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84845
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1816 <STO>
A;Cross-references: UNIPROT:Q9SDB1; GB:AB002093; NID:g6598351; PIDN:AAC02761.2; GSPDB:GN
C;Genetics:
A;Gene: At2g41700
A;Map position: 2

Query Match 16.3%; Score 1813; DB 2; Length 1816;
Best Local Similarity 27.0%; Pred. No. 4.3e-105;
Matches 579; Conservative 287; Mismatches 633; Indels 646; Gaps 70;

QY 260 DSSLSPACSEL-----IGALDHPHLSRLWRLKPLILGK---LLFAPDPTFTRKMA 309
DB 12 DTTTHPAHSNIDKTVVEVGKGNSSPPEVL-----KLLAEGDFLAFAPDPTDETNMID 66
QY 310 QVNRTEBELLLRDVREVMELGPRIFFTFNDSNVAMQLQLQMDGRRQRPRGGRDH 369
DB 67 ILSLKPPELRLV-----TKIF-----KDD 85
QY 370 MEALRSFLDPGSGYGMQDAHDVGHVLTGLRVTCLSLDKLEAPSEAALVSRLQLL 429
DB 86 IE-LETVI-----TSAH-----YGVSEVRNC-SNPKIGA-----114
QY 430 AEHRFWAGVVF--LGPE--DSSDPTHTPDGLGPHVRIKIRMDIDVVTNKKIRDFWD 485
DB 115 -----VVFHEGPHLFDYSIRLNHTW--AFAPGPNVKSIND-----TN-----150
QY 486 PGP-----AADPLDRLYVWGCVYLQDLVERAAVRV-----LSGANPRAGI---527
DB 151 -GPYINDLEMGINTIPTMQYSFSGFLTLQOVVDSFIIFASQONNDLPLSHNSLSALRFE 209
QY 528 -----YLOQMPYPC--YVDDVFLVLSRLPLFLTLAWIYSVTLTKAVVREKETR 576
DB 210 LPWTLFSPSVIRWVPPTREYTDDEFOSIVKSVNGLEF-----247
QY 577 LRDTMRAMGLSRVLMGLWFLSCGLPFLLSAALLVLVLKLGDLIPYSHGVPVFLAFA 636
DB 248 -----PKYSDKTLVFTYFFLFG 264
QY 637 VATVTSQFLLSAFPSRANLAACGLAYFSLYLPY-----VLCVARNRDLPAGR 686
DB 265 LSAILMSFMITSTPFRTRAKTAVAVGTLFLGAFPPYTYNDESVMWL-----K 312
QY 687 VAASLLSPVAFGFCESLALLEQGAQHNVGTRPTADVSLAQVSGLLLLDAALYGL 746
DB 313 VVASLLSPTAFALGSIINFADYERAHVGLRWSNIWRASSGVFFVCLL--MWLLDSILYCA 370
QY 747 ATWYLEAVCPGQYIGIPFPNPPFRRSYW-----CGPRPPKSPACPT-----PLDP 792
DB 371 LGLYLDKVLPRENGVRVPMNPFIPSKYFGRKKNLQNRIPGFTDMFPADIEVNGQEPDP 430
QY 793 ---KULVEEAPPLGSPG-VSVRSLEKPPSPQP--ALRGLSLDPYOGHITAFIGHNGAG 846
DB 431 VPESISLEMRQOQLDGRCIQVRNLHKYVYASRRGNCCAVNSLQLTLYENQILSLIGHNGAG 490
QY 847 KTTTSLISGLFPSPGSAFTLGHVRSMAAIRPHLGVCPCQVNVFLDMLTVDHVVWFY 906
DB 491 KSTTISMLVGLLPTPSGDALILGNSITNMDIEIKELGVCPQHDLPELTVREHLEWFA 550
```

```
QY 907 RLKGLSAAVVGPEQDRLLQDVGIVKSKSVQVTRHLSGEMOKLSVAIAFVGGSQVILDEP 966
DB 551 VLKGVBEGLSKSTVDMABEVLSDKINTLVRALSCKMKRLSLGIALIGNSKVILDEP 610
QY 967 TAGVDPASRRGIWELLKTKYREGRTLLSTHLDDEAELLGDRVAVVAGGRCCGSGPLFLR 1026
DB 611 TSGMDPYSNRLTWOLIKIKKGRILLTTHSMDAEELGDRIGIMANGSLKCCGSSIFLK 670
QY 1027 RHLGSGYIYTLVKARLPLTTNEKADTDMEGSVDTROEKKNGSGSRVGTPLLALVOHWV 1086
DB 671 HHVGVGYTLTLVK-----TSPTVSAAH-----IVRHI 699
QY 1087 PGARLYEELPHELVLVLPYTGADHSGFATLRELDTRLAEL-----RLTGYGSDTSLEE 1141
DB 700 PSATCVSEVGNESISFKPLPL--ASLPCFENNFRIESCCKNSDDYPGIQSYGYSVITLEE 757
QY 1142 IFLKV-----VBECAADTDMEDSGCOHLCTGIAGLDVTLRLKMPPOETALENGEPAG- 1194
DB 758 VFLRVAGCNLDIEDKQEDIFVSPDTKSLVCIG-----SNQKSSMQPKLLASCN-DGAGV 811
QY 1195 ---SAPE-----TDGSGPDAVGRVQGWALTRQQLQALLKFLARRS 1235
DB 812 IITSVAKAFRLIVAAVMTLIGFISIQCCGCSIISRSNFW---RHCKALFIKARSACRD 867
QY 1236 RRGCPAQIVLPALFVGLALVFSLIVPPGHYPALRLSPMYGAQVFFESDAPGDPGRAR 1295
DB 868 RKTVAQFQIIPAVFLFLGLLF-LQLKPHDPQKSIIT------TAYFNLLSGKGGG- 918
QY 1296 LLEALLQEALEBPVQVHSHRSFAPEVAEVAKLASGNWTPESPSPACQSQPGARRL 1355
DB 919 -----PIPFD---LSVP-IAKEVAQYI-EGGWI--QPLRNTSYKFPNKEA 957
QY 1356 LPDCAAAAG---GPPPPQAVTSGEVVQNTGNRLSDFLVKTYPRLV---ROGLKTKKWV 1409
DB 958 LADAIDAAGTLP------TLLSMSEFLMSSPDQSQSSREGLSHSDSC 1001
QY 1410 NEVRYGGFSLGGDPGLPSQELGRSVEELWALLSLPLPGGALD-RVLKNLTAWAHSLDAQ 1468
DB 1002 NH-----PDGSLGYTLVHNGTC-----1018
QY 1469 DSLKIWNKNGHSMVAFVNRASNAIIRALPPLPGPARHAHSITTLNHPNLNLTKEQ-LSEA 1527
DB 1019 -----QIAGPIYINVMAAILRL-----ATGNKNWTIQRNHPPLPTQRIQRH 1063
QY 1528 ALMASSVDVLVSTCVVFMAMSFVPASFTLVLIBERVTRAKHLQMLGGLSPTLYLWLNFLWD 1587
DB 1064 DLDAFAAAIIVNI---AFSFI PASFAVPIVKEREKAKHQOLISGVSVLSYMLSTTVWD 1119
QY 1588 MCNVLVPACIVLILFLAQORAVYAPANLPALLLLLLYGVWSITPLMYPASFPSPVSTA 1647
DB 1120 FISFLFPSTPAIILFYAFGLEQFIGIGRFTPLVLMLEYGIAIASSTYCLTFFPTEHMA 1179
QY 1648 ---YVVLTCINLFIGINGSMATFVLELPSDOKLQEVSR1-----1683
DB 1180 QATSSSVLLPI SILVFPSSNVLVWVHFFSGLILAWISFVWGLIPATASANSYLKELL 1239
QY 1684 ---LKOVFLIFPHCLGRGLIDMVRNQAMADAFERLG--DROFQSPRLREVVGNK----1733
DB 1240 FRVALQNFRFLSPGFCFCDGLASL-----ALLRQGMKDKSHGVFEWNVGTGASTCYL 1291
QY 1734 -----LLAMVIOG---PLFLFTLLLOHRSQLLPQPRVRS1-----PLLG--1770
DB 1292 GLEVRLEYCYSMLLLSFFHIGIDTKUSLIYTGASRLTELI-YDRVYSTSFSTFPLKDS 1350
QY 1771 -----EDEDVARERVRVQCATOGDVLVLRNLTKVYRGQR---MPAVDRCLGIP 1818
DB 1351 TGAISTMEDDIDVOEERDVI SGLSDNTMYLQNLKRVYPGDKHGPKVAVQSLTFSVQ 1410
QY 1819 PGSCFLLGLVNGAGKTTSTFRMTGDTLASGEAVLAGHSVAREPSAAHLSMGYCPQSDAI 1878
DB 1411 AGEFCFGLTNGAGKTTLSMLSGEETPTSGTAFIFGKDIVASPKAIRQHIYGCYCPQFDAL 1470
```

Qy 1879 PELITGHEHLELLARLGVPEAQAQTAGSGLARLGLSWADRPAQTYSGNKRKULATAL 1938
Db 1471 FEYTLVKEHLELYARIKGVVDHRIIDNVVTEKLVDFLLKSHKPSFTLSGNKRLSLVAI 1530
Qy 1939 ALVGDPAAVFLDEPTTGMDPSARFLWNSLLAV-VREGR-SVMLTSHSMEECALCSRLA 1996
Db 1531 AMIGDPPVILDEFTGMDPVAKGFMDVTSRLSTRSGKTAVILTTHSMNAQALCTRIG 1590
Qy 1997 IMVNGRFRCLGSPQHLKGRFAAGHTLTLRVP-----AA 2029
Db 1591 IMVGGRLRCISPOHLKTRY--GNHLEVPFYGVKPEVNSVLENFCQIIQQWMLFNV 1648
Qy 2030 RSQPA----- 2035
Db 1649 PTQPRSLGDLLEVCIGVSDSITPDTASASELSUSPEMVQRIAPLGNQEVSTLVPLPE 1708
Qy 2036 -----AFVAAEFPGSELREAHGGRLR 2057
Db 1709 EDVRFDQLSEQLFRDGGIPLPIPAEWLTKFESALDSFIQSSFFGATPKSCNGLSIKY 1768
Qy 2058 QLPFG-CRCALARVFGELAVHGAHGVEDSVQTMLEEVFLYPS 2101
Db 1769 QLPFEGGLSLADAFGHLERNRNLRGIAEYSISQSTLETIFNHPA 1813

RESULT 8
C88925
Protein P33E11.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C88925
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/c.elegans/ and www.sanger.ac.uk/projects/c.elegans/ and www.sanger.ac.uk/projects/c.elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C88925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1317 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC17542.1; PID:g3158495; GSPDB:GN00023; CESP:F33E11.4
C:Genetics:
A:Gene: F33E11.4
A:Map position: 5

QY	972	PASRRGIWELLKYREGRTIILSHHIDAEALLGDRVAVVAGGRCCGSPFLRRHLGS	103
DB	342	PGARHETWTLIQREKERTILLATHFMEADLLGDRIAIMAHGQLECCGSPMFLKQOYGD	401
QY	1032	GYLTLVYKARLPLTTNEKADTMDEGSDVTRQEKKGSGSQRVGTPLLALVQHVWPGARL	1091
DB	402	GYHLTIV-----YDTSTDPVSKTWD-----TIREYIPEAHV	433
QY	1092	VEELPHELVLVLPYGAHDGFSFATLFRELDTRLAELRLTCYIGSDTSLEBIFLKVBECA	1151
DB	434	PSYIGQBATYLLSAT--HRPIFPKLFKELEDHOTQCGITSGFVSIITWEEVFLKVGHATD	491
QY	1152	ADTMDMGSCQHLLCTGIAGLDVTLRLKMPPOBTALENGEPAGSAPETDQSGPD--AVG	1209
DB	492	ERVNYEHG-----IEN--DISEMIEKDOPILQDLRAQV	522
QY	1210	RVQGWALTROQLQALLKRFLLARRSRGILFPAQIVLPALFVGLALVPSLIVPPGHYPAL	1269
DB	523	RVTGFTLMOHAKAMFYKRAIFPRKWTQFLPOLVPVAVLVMVFTSVLP-----	574
QY	1270	RLSPTMYGAQVPSSEDAPGCRGLLEALLQEAGLEBPPVQSHSRPSAPVPAEVAK	1329
DB	575	-----SVKE-----	578
QY	1330	VLASGNWTPESPACQSQOPGARRLLPDCPAAAGPPPPQAVTGSVEVVQNLTG--RNLS	1388
DB	579	-----QDPQTISLAPPSDTKKAG-----HLVSDSGNYVTLGGSQNLS	616
QY	1389	DFLVKTYPRLVROGLTKKWNVRYGGFSLGGRDPLSGQBLGRSVEELWALLSPLPG	1448
DB	617	SMV-----QGTVTQLGVTO-----TVVDITSNVKE-----	641
QY	1449	GALDRVLKNTLAWAHSLDAQDSLKIWFNNKGHSMVAFVNRASNAILRAHLPGPARHAH	1508
DB	642	-----FMDQTNAM-----GS	652
QY	1509	SITTLNHLNLTKEQLSEALMASSVDVLVSIQVWFAMSPVPASFVLVIEEVRTTRACHL	1568
DB	653	RTFGLHYALGFVSPMFNFSTVPSLKISPAVCV-----AGYSOFLITERKKSKGM	704
QY	1569	QLMGGLSPTYLWGNFLWDMCNVLVPACIVLVLFLAFOQRAYVAPANLPALELL--LLLYG	1627
DB	705	QLLSGIRPMFWLTAFTDAEAWFVIRILCPDAIFYIFENITAYTHDFGVMLTLTSLFLLYG	764
QY	1628	WSITELMPASFEFSPSTAYVVLTCINLFIGINGSMATFVLE-----	1670
DB	765	WTALPTFYEQPFESAPKGFMMVTHILTMIGSIAVPIISQTSLSLDAGYUWSIIFAW	824
QY	1671	LFSDQKLQEVSR1-----LKQVFLIFPHF-----CLGRGLDMVRNQAMADAFER	1715
DB	825	LFTYNIQSIAVTTFQENVRVIACKKLDCTIPMKVATACCG-----TASER	871
QY	1716	L-----GDRQFQSPLRWEVVGKNLLAWI-----QGFLFLTLLQHR--SOLLPOQ	1761
DB	872	LYVDNVLFVGNR-----KGILVYVFLAVQGFYIWIWVFWPMEHQDTKLFALI	919
QY	1762	RVRSL--PLLGEDEEDVARERE-----RVQOGATQGD--VLAVRLNTLKVYQGQM	1807
DB	920	RCKRADNPIDWIITDTDKVDERVEDSDVIAKBSWQORLANNKNTALVSNLWKVY--GNF	977
QY	1808	PAYDRLCGIPPCGCLLGVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAREPSAAHL	1867
DB	978	NAVKGYNFVNSKDCFLGLGVNGAGKTSITFQMLTGENSISSGDAYNGHSVKQNWREAGA	1037
QY	1868	SMGYCPOSDAIPELLTGREHLELLAHLRGVPEAQVAQTAGSGLARLGLSWYADRPAGTYS	1927
DB	1038	NTGYCPOYDAIIXEMSGEETLYMFARIRGIPEKDIPKYNVAHIHIGIWNYSASRQIKTYS	1097
QY	1928	GGNKRKLATALAVGDPVAVVLDPTTGMDDPSNRRFLNLSLLAVVREGSRVMLTSHMEE	1987
DB	1098	GGNKRLSGLIAVGLPVDLLDEPTSGVDPKARRIINWLRDLRGDALVLTSHSMDE	1157

Cy 1988 CEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAR - SQPAAAFVAEPFGSE 2046
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 1158 CEALCELTAIMVYGFRCYGSCQHISRYSGVTLLIRLNKNRNDAEKTKSTIKQTFRGSV 1217
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Cy 2047 LREAHGGRLRFOLPPGCCALARYFVGELAVHGAEHGVDEFSVSOTMLEEVFLYFSKDQG 2106
:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1218 IKEHVVLQLNFDIPRDGD - SWSRLFEXLETVSTSLNWDDYSLSQTILLEQVFIEFSRDAGV 1276
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Cy 2107 DEDTEEQEAGVGVDPAAPGLQH 2128
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1277 SSDSEFDGASTGSANSRGQN 1298
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
T15200
hypothetical protein F12B6.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
C/Accession: T15200
R/Pauley, A.; Maggi, L.
submitted to the EMBL Data Library, May 1997
A>Description: The sequence of C. elegans cosmid F12B6.
A.Reference number: Z18307
A.Accession: T15200
A.Status: preliminary; translated from GB/EMBL/DDDBJ
A.Molecule type: DNA
A.Residues: 1-1447 <PAU>
A.Cross-references: EMBL, AF003138; NID: g2088708; PID: AAB54153.1; GSPDB: GN
A.Experimental source: strain Bristol N2; clone F12B6
C/Genetics:
A/Gene: CESP: F12B6.1
A/Map position: 1
A/Introns: 79/2; 114/3; 177/1; 224/3; 331/1; 345/3; 373/2; 417/2; 464/1; 536/1; 659/2; 6
C/Superfamily: ATP-binding cassette homology

Query Match 14.4%; Score 1600.5; DB 2; Length 1447;
Best Local Similarity 27.3%; Pred. No. 6.8e-92;
Matches 475; Conservative 257; Mismatches 533; Indels 477; Gaps 51;

Cy 492 PLTDLRVVWGCFVYLQ-----DLVERAAVRVLG--ANPRAGLYLOOMPY 534
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 13 PLIDLKYITFGSFLQSWSPLESKSQKLSESIDRAIMSLELTNQDTANGLVYAQQEYP 72
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Cy 535 PCYVDDVFLRVLRSPLFLTAMIYSVTLTVKAIVREKETRIERDTRMAMGLGRAVLWLIG 594
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 73 PCTVKDFNFVAL--FMPLFLLSIFFSALLNVKVIVEQEKIKEQWRAMLGDVAHFIS 130
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Cy 595 WFLSCLGPFLLSAALLVLVLKLGDIILSYSHGVVFLFLAAPAVATVTSQSFLLSAFTSRAN 654
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 131 WGLISLVLFNISVIIISIISKVXIFDYTDYTLILLFVLIILFPSSIAMSIFFSTLETFNAN 190
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 655 LAACGGGLAYPSLYLPVYLCVANRDLRPAGGRVAASILLSPVAFPGCESIALLEEQGEA 714
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 191 IATAATCVLWFVPFIPTQLLTRDISPTFNRI SL- ILPTAMGH-CFKLLESFNAMERA 248
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 715 QWHNVG--TREPDVFSLAQVSGELLDAALYGLATWTLEAVCPGOYGIPEPNWFPFRS 772
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 249 TWSDIWNMNVPVIG-SVELCMIMLVVDVAVFLILAWYISAVAPGDFGVQPWFPTLK 307
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 773 YWCGRPPKSPACPFTPLDPKVLVEEAPPGLSGVSVRSLE----KFPGSP-----Q 821
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 308 YW-----APGLYNRVFEVDHEFTIPNSDSFDS 338
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 822 P-----ALRGLSLDFYQCHITAFTHGNAGAKTTTILSLGFLPPSGSAGFILGHDVRRSM 876
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 339 PTNLTLALDCINLRLYESGITGLIGHNGAGKTMTTMSILCGLYAPSSTGTAKIQRIDNTDL 398
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 877 AAIRPHLGVCPOYNNVLPMDLTVDBHWFWFYGRCLKLSAAVVGPEODRLLDQVLGVKSQSVQ 936
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db 399 RRVRDVLGICQHNVLFSHLTIVSEQRLPALAKGVDPSEILT SQVDELIASVSLTEKAN KL 458
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:
Cy 937 TRHLSGGMQRKLSVAIAFVSGSVQVVIDEPTAGVDPASRRGIWEMLLKKYREGRTLILSTH 996
|||:|||::|||::|||::|||::|||::|||::|||::|||::|||:

Db 1443 -GRFTAVNELCLA VDQKECFGLLVNGAGKTTTNTILTGOSFASGEAMIGGRDVTB- 1498

Qy 1864 AAHL5MGYCQSDAI FELLTGRHEHLELLARLGVPEAQVACTAGSLARLGLSWYADRP 1923

Db 1499 --LISIGYCQFDALMDLDTAGRESLETLAQHGCFENYKA--KAEILILECVGMIAHADKLV 1554

Qy 1924 GTYSGGNKRKLATALALVGPVAVFLDEPTTGMDPSARRFLWNSLLAVVREGS-VMLTS 1982

Db 1555 RFYSGGQKRKISVGVALLAQTOMIILDEPTAGIDPKARREVWELLWCRESHSALMLTS 1614

Qy 1983 HSMEECEALCSRLAIVMNGFRCLGSPHLKGRFAAGHTILTLRV--PAARSQPAFAAVAA 2040

Db 1615 HSMDECEALCSRLAVLNRSGLIALGSSQELKSLGYNNYTWLTSIYENQORDM-VVOLVOT 1673

Qy 2041 EPPGSELRHAGGR--LRFQLPPGRCALARVFGELAVHGAHBGVDFSVQSMLEEVF 2097

Db 1674 RLPSVLKTTSTNKLNLKQIPEKEDCWSAKFEMVQALAKDLGVKDFILAQSSLEETF 1733

Qy 2098 LYFSKQDQKDED 2109

Db 1734 LRLA---GLDED 1742

RESULT 11

S60124

transport protein homolog C48B4.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 13-Jan-1996 #sequence_revision 12-Apr-1996 #text_change 16-Aug-2004

C;Accession: S60124; S40724; S40725

R;Kershaw, J.

submitted to the EMBL Data Library, November 1995

A;Reference number: S60124

A;Accession: S60124

A;Molecule type: DNA

A;Residues: 1-1767 <KER>

A;Cross-references: EMBL: Z29117; NID: g439247; PID: g1066912

C;Genetics:

A;Map position: III

A;Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/

C;Superfamily: ATP-binding cassette homology

C;Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein

F;628-818/Domain: ATP-binding cassette homology <ABC1>

F;645-652/Region: nucleotide-binding motif A (P-loop)

F;764-769/Region: nucleotide-binding motif B

F;1457-1642/Domain: ATP-binding cassette homology <ABC2>

F;1474-1481/Region: nucleotide-binding motif A (P-loop)

F;1586-1591/Region: nucleotide-binding motif B

Query Match 12.6%; Score 1407; DB 2; Length 1767;

Best Local Similarity 25.9%; Pred. No. 1.3e-79;

Matches 485; Conservative 298; Mismatches 665; Indels 424; Gaps 59;

Qy 383 GYSQDAHA--DVGHVGLTGRVTECLSD---KLEAPGEAALVSR-ALQLIAE----- 431

Db 159 GYTSKDAAAKTVDDLMKKAERFQSAKLKLVNSESSEQLLTVLRLNPLMNETFC 218

Qy 432 HRFWAGVVF-----LG---PEDSDPTHEPTDPLGPHVRKIRMDIDVVT 474

Db 219 NSYAAAGVVDVDTNKKLVNRIILGKTPETWHLTETSYNPGPSSGR----- 267

Qy 475 RTNKIRDFWDPGPAADPLTLRVYWGCVFYLODLVERAAVRVLSGANPRAGLVLOOMP 534

Db 268 -----YSRIPSSPP-----YWSAFLTFQHAIESFLSSVQSGAPDLPITRLGPE 313

Qy 535 PCYVDDVFLVLSRLPLFLTLAMIYSVTLTVKAVRE---KEPRLDTRMAGLSRAVL 591

Db 314 PRYKTSVSFAID-----FFPFIWAFVTFINVIHITREIAENHVAVKPYLTAMGLSTFMF 368

Qy 592 WLGHFLSCGLGFLLSAALLVLVLKGLDILPVSHGVGVFLFLAAFAVATVQSFLLSAPFS 651

Db 369 YAAHVWMAFLKFFV--IFLCSIIPLTFMFEVSPVPAALIVTVLMYGLGAVIFGAFVASFFN 426

Qy 652 RANLAAACGGGLAYSLVLPYVLCVAVMRDL-----PAGGRVAASLLSPV-----AFGFG 700

Db 427 NTNSAIK-----AIIIVANGAMIGISKLRPELDQISSCFLYGLNINGAFALA 473

Qy 701 CESLA-----LLEQEGEAGQHNHVTTRPTADVFLSAQVS-GLLLLD-----AALYG 745

Db 474 VEAISDYMRREERELNTNMFNDSLSLHFLSGWALVMMIVDILMWSIGALVVDHIRTSDAFS 533

Qy 746 LATWY-----LEAVCPGOYGIPEPNWPFPRRSYWCWGPSPKSPAPCPTPLDPKV 794

Db 534 LRTLDFEAPEDDENOTDGTVAQTRINEQVRNRVRSD-MEMNPMASTSLNPNADSDS 592

Qy 795 LVE-----EAPGLSPGVSRSLERKFPSPALRGLSLDFYQGHITAFLEHNGAGKT 848

Db 593 LLEGSTADGARTADAIIVRNVLKWTSTTGERAVDGLSLRAVRGQCSILLGHNGAGKS 652

Qy 849 TTLSIISGLPPPGGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFMDLTVDHWFVYGRLL 908

Db 653 TTFSSIAGIIRPTNGRITICGYDVGNPEGETRRHIGMCPQYNPLYDQLTVSEHLKLYVGL 712

Qy 909 KGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGMOKLSVAIAFVGGSOVVTLDDEPTA 968

Db 713 KGAREKDFKODMKRLSLSDVKLDFKENEKAYNLGGMKRKLVCVCMALTIGDSEVLLDDEPTA 772

Qy 969 GVDPASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGGRGCCGSPFLFRRH 1028

Db 773 GMDPGARQDVQKLVREKANKRTILLTHYMDAERLGDWVFMHSGKLVASGTINQYLKQK 832

Qy 1029 LGSGYLTLVKARLPLTTNEKADTMEGSVDTFOEKKNQSGQSRVGTPTQLLALV-----Q 1083

Db 833 FGTGYLLTVV-----LDHNGDKRK-----MAVILTDVCT 861

Qy 1084 HWYPGARLVEELPHELVLVLPYTGADHGSFATFRELD-----TR 1123

Db 862 HYVKEAERGEMHQOQIEIILP--EARKKEFVPLFOALEAQDRNRSNVFDNMNTLKSQ 919

Qy 1124 LAELRLTGYSISDTSLEEIEFLKVVECAADTMEDGSCGQHLCTGTAGLDVTLRLKMPQP 1183

Db 920 LATLEMRSGLSLNTLEQVFIIT-----GDKVDKAIASRQ-NSRISHNSR 963

Qy 1184 ETALENGEPAG-----SAPETD-----QGGSPDAVGRVQGWALTRQQLALLKRF 1230

Db 964 NASEPSLKAGYDTQSTKSADSYQKLMDSQARGPEK-----SGVAKMVAQFISIMRKKFL 1019

Qy 1231 LABRRRGLEAQIVLPALFVGLALVFSLLIYPPGHPALRLSPMYAQVVSFFSEADAPGD 1290

Db 1020 YSRNWAQLFTQVLIPIILGLV-----GSLTTLKSNNT-----D 1054

Qy 1291 PGRARLLEALLQAGLEPPVQHSRFSAPVPAEVAKVL-ASG-----NWTPESPSA 1344

Db 1055 QFRS-----LTPSGIEPSKVVRFPENGTTPEEAAKFEKILRKSQGGFEVLNNTKNP--- 1105

Qy 1345 CQCSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQNLNLTGRLNLSDFLVKTYPRLVROGLK 1404

Db 1106 -----LPNITKSLIGEMPATI----- 1122

Qy 1405 TKWVNEVRVYGGFSLGGRDFGLPSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHS 1464

Db 1123 -----GMTMN----- 1127

Qy 1465 LDAQDSLKIWFNNKGWH---SMVAFVNRAS-NATLRAHLPPGPARHAHSITTLNHLNLT 1520

Db 1128 ---SDNLEALFNWRYHVFLTSLMINARLCTVDAEISSGVFLYSKSTSN-----SNLL 1180

Qy 1521 KEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVLIIEERVTRAKHLQMLGGLSPLYW 1580

Db 1181 PSQI-----IDVLLAPMLIIFAMVTSTFMVLEIBERTCQFAHQOQLTGISPTIFY 1231

Qy 1581 LGNFLDMCNLYLPACTVLIIFLAFQORAVAPANLPALLLLLLLYGWSITELMYPASFF 1640

Db 1232 SASLIYDGILYSILCLIFLFWELAF-----HWMYDHLAIVILFWFLYFSSVPFIYANSFL 1287

Qy 1641 FSVFSTAYVVLTCINLFIGINGSMATFVLELFSFDQKLQE-VSRILKOVFL-IFPHFCLGR 1698

Db 1288 FQSPSKANVLL--IIQVVVISGALLAVFLIFMIFNIDWLKSLILNIFMFLPSYAFGS 1345
Qy 1699 GLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKILLAMVIOGPL--FLIFTL----- 1750
Db 1346 AII-TINTYGMILPSEEL-----MNDHCGKNALWMTGVCFSFALFVLLOPKFYR 1396
Qy 1751 -----IQHRSQLLPQPRVRSPLPLGGEDEVDARERERVQAGTGDVLVLRNLTKVYR 1803
Db 1397 FLSQVMTVRSSHNVPMMGDLFVC-----ESVSEERERVHRVNSQNSALVIKDLTKTF- 1451
Qy 1804 GORMPAVDRLCIGIPGECFGLLVNGVNGACKTSTFRMVTGDTLASGEAVLAGHSVAREPS 1863
Db 1452 -CRFTAVNELCLAVDOKECFGLLVNGVNGACKTTFNLTGQSFSSGEAMIGGRDVT----- 1507
Qy 1864 AAHLGMYCQPSDAIFELITGREHLELLARLRGVPFAQVAQTAGSGLARLGLSWADRP 1923
Db 1508 --LISIGYCPQFDALMDLTGRESLEILAQMHPENYKA--KAEILBECVGMIAHADKLV 1563
Qy 1924 GYISGKRRKLTALALVGDPAVFLDEPTTGMDPSARRFLNLSLAVVRBGRS-VMLTS 1982
Db 1564 RFYSGGQKRKISVGVALLAFTQMIILDEPTAGIDPKARREVWELLIMCREHSNSALMLTS 1623
Qy 1983 HSMESCEALCSRLAMVNGFRCLGSPQHLKGRFAAGHTLTLRV--PAARSQPAFAFVNA 2040
Db 1624 HSMDSCEALCSRLAVLNRLSLTAIGSSQELKSLYNNYTMTLGSLYEPNQDM--VVQLVOT 1682
Qy 2041 EPPGSELREAHGGR---LRFQLPPGRCALARVFGELAVHGAEGHVEDPFSVQTMLEEVF 2097
Db 1683 RLPNSVLKTTSTNTKLNKWLQIPKEKEDCWSAKFEMVQALAKDLGVKOPILAQSSLETF 1742
Qy 2098 LYSFKDQKQKDED 2109
Db 1743 LRLA---GLDED 1751

RESULT 12
T42749
ATP-binding cassette transport protein homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2004
C:Accession: T42749
R:Wu, Y.C.; Horvitz, H.R.
Cell 93, 951-960, 1998
A:Title: The C. elegans cell corpse engulfment gene ced-7 encodes a protein similar to A
A:Reference number: Z22259; MUID:98297348; PMID:9635425
A:Accession: T42749
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1704 <WUY>
A:Cross-references: UNIPROT:P34358; EMBL:AF049142; NID:g3172340; PIDN:AAC24116.1; PID:g3
A:Note: ced-7
C:Genetics:
C:Superfamily: ATP-binding cassette homolog

Query Match 12.6%; Score 1405; DB 2; Length 1704;
Best Local Similarity 25.9%; Pred. No. 1.6e-79;
Matches 486; Conservative 297; Mismatches 668; Indels 422; Gaps 59;

Qy 393 GYSWQDAHA--DVGHILVGTIGRVTCLSLD---KLEAAPSEALVSR-ALQLLAE----- 431
Db 92 GYTSKDAAKRTVDLDMKKFAERFOSAKLKLVSQKNSSEBQLLTVLRNDLPMLNETFCAI 151
Qy 432 HRFWAGVVF-----LG---PEDSSDPTHTPTDLPDGPCHVIRKIRMDIDVVT 474
Db 152 NSYAGVVFDVVDVTKLNKLVIRLLCKTPEETWHTTETSYNPGSSGR----- 200
Qy 475 RTNKIRDFWDGPAADPLTLRLVYVGVYQLDLVERAAVRVLSGANPRAGLYLQOMPY 534
Db 201 -----YSKIPSSPP-----YWTSAPLTFQHAIESFLSSVQSGAPDLPITLRGLPE 246
Qy 535 PCYVDVDFLRVLSRLPLFLTLAWIYSVTLTVKAVRE---KETRLRTMAMGLSRAVL 591
Db 247 PRYKTSVSAFID-----FPPPIWAFVTFINVIHITREIAAENHAVKPYLTAMGLSTFMF 301

Qy 592 WLGWFLSCICLPFLLSAALLVLVLKGLDILPYGHPGVVFLFLAAFAVATVTSFLLSAFPS 651
Db 302 YAAHVWMAFLKEFV--IFLCSIIPITFVMEFVSPAALIVTVLMYGLGAVIFGAFVASFFN 359
Qy 652 RANLAAACGLAYFSLYLPLYLVCVARDRL-----PAGGRVAASLLSPV-----AFGFG 700
Db 360 NTNSAIK-----AILVAGAMIGISYKLRPELDQISSCFYGLNINGAFALA 406
Qy 701 CESLA-----LLEBQGEQAQHNHVGTEPTADVFLSAQVS-CLLLLD-----AALYG 745
Db 407 VEASIDYMRERRELNTNMFNDSLSLHSGWALVMVMDILMWSIGALVVDHIRTSADES 466
Qy 746 LATWT-----LEAVCPQYGIPEPMFPFRRS--YWCGRPRPKPAPCPTPLDPK 793
Db 467 LRTLDFPEAPEDDENQTDGVTAGTNRINEOVNRVRRSDMEIQMNPMASTSLNPPNADSD 526
Qy 794 VLVV-----EAPPGLSPGVSVRSLEKFPSPQPALRGLSLDFYQGHITAFILGHNGAGK 847
Db 527 SLLEGSTEADGARDTARADIIVRNLVKIWTSTTGERAVDGLSLRAVRGQCSILLGHNGAGK 586
Qy 848 TTTLISLGLPFPSPGSGSFIILGHVYRSSMAAIRPHLGVCPQYNVLFDMLTVDHWHWFYGR 907
Db 587 STTFSSIAGITRPTNGRITICGYDVGNBEGETRRHIGMCPQYNPLYDQLTVSEHLKLVYG 646
Qy 908 LKGLSAAVVVGPQDRLLQDVLVSKQSVQTRHLSGMORKLSVALAFVCGSVVILDEPT 967
Db 647 LKGAREKQKQDMKELLSDVKLDFKENEKANVLSGMRKRLCVCHALIGDSEVVLDEPT 706
Qy 968 AGVDPASRRGIWELLKYREBRTLILSTHLDDEALLGDRVAVVAGRLCCCGSPFLRR 1027
Db 707 AGMDPGARQDVQKLVEREKANRTILLTHYMDAEARLGDWVFMHSGKLVASGTNYLQK 766
Qy 1028 HLGSGYILTYKARLPLTNEKADTDMESVDTRQEKNGSQGSRVGTQQLALV----- 1082
Db 767 KFGTYLLTVV-----LDHNGDKRK-----MAVILTDVC 795
Qy 1083 QHWVPGARLVEELPHELVLVLPYTGAHDGSPATLPRELD-----T 1122
Db 796 THYVKEAERGEMHQOIEIILP--EARKKEFVPLFQALEAIQDRNRSVNFNMPTLKS 853
Qy 1123 RLAEIRLTGYGISDTSLEEIFLKVVECAADTDMEDSGCGHLCGTGIAGLDVTLRKMP 1182
Db 854 QLATLEMRSFGLSLNTLQVFTI-----GDKVDKAIASRQ--NSRISHNS 897
Qy 1183 QETALENGEPAG-----SAPETD-----QCSQPDVAVRVOGHVLTQQLALLKRP 1229
Db 898 RNASEPSLKPAYDTQSTKSADSYQKLMDSQARGPEK-----SGVAKMVAQFISIMRKFP 953
Qy 1230 LLARRRRRLGFAQIVLPALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVFFSEDA 1289
Db 954 LYSRRNWAQLTQVLPIIILGLV-----GSLTTLKSNNT---DQFSVRS----- 995
Qy 1290 DPGRRARLEALLQEBAGLEPPVQVHSHRFSFAPEVPAEVAKVL-ASG-----NWTPEPSP 1343
Db 996 -----LTPSGIBPSKVWVRFENGTTIPEEAANPEKILRKSGGPEVLNNTKNP-- 1042
Qy 1344 ACQCSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQNLGTBNLSDFLVKTYPRLVROGL 1403
Db 1043 -----LPNITKSLIGEMPPATI----- 1059
Qy 1404 KTKKWNVEVRYGGPSLGGDRDPLPSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAH 1463
Db 1060 -----GMTMN----- 1064
Qy 1464 SLDAQDSLKIFWNKGMH---SMVAFVNRAS-NAILRAHLPPGPARHAHSITTLNHLML 1519
Db 1065 ---SDNLEALFNRYHVHLPTLSMINRARLTGTVDAEISSGVFLYSKSTSN---SNL 1116
Qy 1520 TKEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVILIEERVTRAKHLQMGSLPTLY 1579
Db 1117 LPSQL-----IDVLLAPMLILIFAMVTSFTVMFLIEERTCOFAHQQLGTISPTIF 1167

Qy 1580 WLGNFLWDMCNVLPACIVTVLIIFFLAFOORAYAPANLPALLLLLLLYGWSITPLMTYPASF 1639
Db :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1168 YSASLIYDGIYSLICILIFLFMFELAF-----HWMYDHIAIVILEFWLFYFSSVFIVAVSF 1223
Qy 1640 FFSVPSTAYVVLTICINLFIGINGSMATFVLEFSDOKLOE-VSRILKQVPL-IFPHFCUG 1697
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1224 LFOSPSKANVL--IHWVVISGAALLAVFLIMFNIDEWLKSILVINFMFLPSPYAFG 1281
Qy 1698 RGLIDVRNQAMADAFERLGDRQFQSPLRVEVVGKNLLAMVIQGPL-FLLFTILL----- 1750
Db :|:~::~:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1282 SAII-TINTYGMILPSEEL-----MNWDHGCKNAWLNGTGVCSFALFVLQPKFYR 1332
Qy 1751 -----LQHSQLIPQPRVSLPLLGEDEEDVARERERVQOGATQGDVLVURNITKYV 1802
Db :::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1333 RFLSQVMVTWRSSSHNNVQPMWGDLPGVC---ESVSEERERVHRVNSQNSALVTKDLTKTF 1388
Qy 1803 RGOMPFAVDRLCIGIPGECFGILLGVNCAKCTTFRMVTGDTLASGEAVLAGHSVAREP 1862
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1389 --GRFTAIVNELCLAVDOKECFGILLGVNCAKCTTTFNILTQSFPASGEAMIGRDYTE-- 1444
Qy 1863 SAASHLMGYCPQSDATPELLITGREHLELARLRGVPEAOVAQTAGSGIARLGSWVADR 1922
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1445 ---LISIGYCPQDPALMDLTGTRESLEILAQMGFENYKA--KAEILLECVGMIAHADKL 1499
Qy 1923 AGTYSGGNKRLATALALVGDPVAVFLDEPTTGMPRSARRFLWNLSLIAVVREGS-VMLT 1981
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1500 VRFYSGQKKRKISGVALLAPTOMIILDEPTAGIDPKARREVWELLWCREHSNSALMLT 1559
Qy 1982 SHSMEECEALCSRILAMVNGRFRCLSGPQHUKGRPAAGHTTLTRV--PAARSOPAAAFVA 2039
Db ||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1560 SHSMEECEALCSRILAVNRGSLTAIGSSQBLKSLGNNTYMTLSLYEPNQORDM-VVOLVQ 1618
Qy 2040 AEPGSELREAAGR---LRFOLPPGRCALRVFGELA VHGAEGHVDEDFSVSOTMLEEV 2096
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1619 TRLPNSVLKTTSNTKTNLNKAWQPKPEKEDCWSAKFEMVQALAKDGLVKDFILAQSSLEET 1678
Qy 2097 FLYFSKDQDKDED 2109
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1679 FLRLA---GLDED 1688

RESULT 13

T00826
hypothetical protein T32G6.22 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00826
R:Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A:Description: Arabidopsis thaliana chromosome II BAC T32G6 genomic sequence.
A:Reference number: Z14163
A:Accession: T00826
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1246 <R>
A:Cross-references: UNIPROT:Q858T6; EMBL:AC002510; NID:g2618683; PID:g2618705
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 2
A:Introns: 33/3; 95/2; 113/3; 137/3; 168/3; 361/3; 421/2; 432/3; 493/3; 521/3; 535/3; 568/3
A>Note: T32G6.22

Query Match 12.3%; Score 1376; DB 2; Length 1246;
Best Local Similarity 28.7%; Pred. No. 6.7e-78;
Matches 403; Conservative 195; Mismatches 415; Indels 392; Gaps 42;

Qy 927 VGLVSKQSVQTRHLSCGMQRKLSVAIAFVGSGSVQVILDEPTAGVDPASRRGIWELLIKYR 986
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1 VGLSDKINTLVRLSGMKRKLGIALGNISKVILDEPTSGNDPYSMELTWQLIKKI 60
Qy 987 EGRFTLILSTHHDEAELLGDRVAVAGRLCCCGSPFLRRHLGSGYYILTVKARULPTT 1046
Db ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 61 KGRILLTHSMDEAELGDRIGIMANGSKLCGSSSIFLKHYGVGYTLTVK---- 113

Qy	1047	NEKADTMEGSDVTROBKKNQSGSRVGT	POLLALVOHWVPGARLVEELPHELVLVLPYT	1106
Db	114	-----TSPTVSAAH-----	-----IVRHIPSTATCSEVGNESFKUPL-148	
Qy	1107	GAHDSFATIFRELDTRLAEL-----	RLTGVGSDTSLERIFLKV-----VRECAADTD	1155
Db	149	ASLPCENFRETESCMKNSDSDY	PGIQSYGIVTLEEVFLRVACNLDIEDKQBDIF	207
Qy	1156	MEDGSCQHLCTGIAGLDVTLRLKMP	QETALENGEPAG-----SAP-	1198
Db	208	VSPDTKSSLCIG-----	SNQKSSMQPKLLASCN-DGAGVIITSVAKAFRLIVAAVWTLI	261
Qy	1199	---TDQSGDPADVRCVOCWALT	QQLOALLKXPELLARRSRGLFAQIVLPALFVGLALV	1255
Db	262	GFISIQCCGCSIIISRNFW---	RHKALFPIKARSACRDKTVAFQIIPAVFLFLGLL	317
Qy	1256	FSLIVPPGHVPAIRLSPTMYGAQVS	FFSDAPDGPGRARLLLEALLQEAAGLEPPVQHS	1315
Db	318	F-LQLKHPHPOKSIITL---	---TAYFNLLSGKGGG-----PIFPD-	355
Qy	1316	HRSFAPVPAEVAKVLAGSNWTPES	PSPCAQCQOPGARRLLLPCDPAAG---GPPPPQAV	1372
Db	354	--LSPV-IAKEVAQYI-EGGWI-	QPLRNTSYKFPNPKALADADAAGPTLGP-----	401
Qy	1373	TGSEVVONTGRNLSDFLVKTYPRLV-	--ROGLTKKWNVEYRGFSLGDRDPGLPSG	1429
Db	402	-----TLLSMSEFLMSSDQVS	QSGRESLSDSCNH-----	433
Qy	1430	QELGRSVEELWALLSPUGGALD-	RVLKNLTAWAHSLDAQDSLKIWFNNKGWHSMAFVN	1488
Db	434	-----PDGSLGYTVLHNGTC-	-----QAGPTIYIN	457
Qy	1489	RASNAIIRAHLPDGPBARHANS	IITLHPLNLTKEQ-LSEAAIMASSVDVLVSICVWFMS	1547
Db	458	VMHAAILRL-----	ATGNKNMTIOTRNHPLPPTKTQRIQRHDLDAFSAIIIVNI-----AFS	509
Qy	1548	FVPASFTLVIIIEERVTRAKHLQ	LMGGLSPTLYWLGNEFLWDMCNVLPACIIVWLIFLAFOQ	1607
Db	510	FIPASFAVPVIVEREVAKHQ	LLISGVSVLSTYVMDFISFLPSTFAIILFYAAGL	569
Qy	1608	RAYVAPANLPALLLLLLYGWS	ITPLMYPASFPFSPVSTA-----YVVLTCINLFIGING	1662
Db	570	EQFTIGRFLPTVLMLEYG	LTAISSTYCLTFFTFTHESMAQATSSYGVLLPISLVSFSS	629
Qy	1663	SMATFVLELPSDQKLQSVRI-	-----LKQVFLIFPHFCLGRG	1699
Db	630	SNVILMVHFFSGLLILMWIS	FWNGLIIPATASANYIKELIIFRYALQNFRLSPGFCFSDG	689
Qy	1700	LIDMVRNQAMADA	FERLG---DROFQSPRLWEVVGKN-----LLAMVIQ	1740
Db	690	LASL-----	ALLRQGMKDKSHGVFENNVTGASICYLGLVLRVLCRYSMILLSPFH	741
Qy	1741	G---PFLFLTLLQHRSQLLPQ	VRSL-----PLLG-----EDEDVARRER	1782
Db	742	GIDTKUSLITYTIGASRLTELI-	YDRVYTSFSTFTEPLLKDSTGAISTMEDDIDVQEBDR	800
Qy	1783	VVQATQGDVLVRLNLTKVYRGQR-	---MPADVRLCLGIPPGCEGFLGNGVNGAKTSTFR	1838
Db	801	VISGLSDNTMLYLQNLKRY	PGDKHGPVAVQSLTFESVQAGECFGFLGTNGACKTTLS	860
Qy	1839	MVTGDTLASGEAVLAGHSVAR	BPSAHLBMGYCQSDAIFELLTGREHLELLARLQVP	1898
Db	861	MLSGEETPTSTAFIFGKD	IVASPKAIRHIGYCPQFDALFEYLTVKEHLEYARIKGV	920
Qy	1899	EAOVAQTAGSGLARLGLSWADR	PAGTYSGNKKKLATALLVGDPAVVFDEPTTGMDP	1956
Db	921	DHRIDINVUTEKLVFEDLLK	HKHKSPFTLSGNGKKKLAVATAMIGDPFIVILDESTGMDP	980
Qy	1959	SARRFLWNSLIIV-VREGR-	SVMLTSHSBECEALCSRLAIMVNGRFCLGSPHLLKGRF	2016
Qy	981	VAKFEMDVTSLSTSTRSGTAV	ILTHSHMNEAOALCTRIQIMVGRRLRCIGSPHLLKTRY	1040

```
QY 2017 AAGHTLTLRVP-----AARQPAA----- 2035
Db 1041 --GNHLEVPFYGVKPNVSNVLENFCCIQQWLFNVPTQPSRLGDLVCIGVSDS 1098
QY 2036 ----- 2035
Db 1099 ITPDTASASEISLSPMWQRIAKFLGNEQVSTLVPPLPEEDVRFDQDQSEQLFRDGGIP 1158
QY 2036 -----AFVAAEPGSELRAHGGRRLRFQPPG-GRCALARVFGELAV 2076
Db 1159 LPFAEWMLTKFKFALUSFIOSSPFGATPKSCNGLSIKYQLPFGEGGLSLADAFGLER 1218
QY 2077 HGAHGVDFSVQTMLEVFYFS 2101
Db 1219 NNRRLGIAEYSISQSTLETIFNHA 1243

RESULT 14
T46467
hypothetical protein DKFZp434E1030.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Aug-2004
C:Accession: T46467
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T46467
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-269 <AAA>
A:Cross-references: UNIPROT:Q9NT80; EMBL:AL1137481
A:Experimental source: adult testis; clone DKFZp434E1030
C:Genetics:
A>Note: DKFZp434E1030.1
C:Superfamily: ATP-binding cassette homology

Query Match 12.1%; Score 1345; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 6.1e-77;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1720 QFQSPLEWVGKLLAMVIOGFLFTLLQHRSQLLPQVRVSLPLGSEEDVARE 1779
Db 1 QFQSPLEWVGKLLAMVIOGFLFTLLQHRSQLLPQVRVSLPLGSEEDVARE 60

QY 1780 RERVVGATQGDVLRLNLTQVYRGQMPADVRLCIGIPGCEFCGLLVGAGKSTFRM 1839
Db 61 RERVVGATQGDVLRLNLTQVYRGQMPADVRLCIGIPGCEFCGLLVGAGKSTFRM 120

QY 1840 VTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPE 1899
Db 121 VTGDTLASRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPE 180

QY 1900 AQAQTAGSLARGLSWADRPAGTYGGNKRKLTALALVGDPAVVFLEPTTGMDPS 1959
Db 181 AQAQTAGSLARGLSWADRPAGTYGGNKRKLTALALVGDPAVVFLEPTTGMDPS 240

QY 1960 ARRELWNSLLAVVREGSVMLTSH 1983
Db 241 ARRELWNSLLAVVREGSVMLTSH 264

RESULT 15
T27121
hypothetical protein Y53C10A.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004
C:Accession: T27121
R:White, S.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z20314
A:Accession: T27121
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
```

```
A:Residues: 1-1564 <WIL>
A:Cross-references: UNIPROT:Q9XM49; EMBL:AL033536; PIDN:CAA22142.1; CESP:Y53C10A.9
A:Experimental source: clone Y53C10A
C:Genetics:
A:Gene: CESP:Y53C10A.9
A:Introns: 43/3; 92/2; 148/2; 226/3; 354/1; 712/3; 817/1; 875/1; 916/3; 984/3; 1069/2; 1
C:Superfamily: ATP-binding cassette homology
```

```
Query Match 9.3%; Score 1038.5; DB 2; Length 1564;
Best Local Similarity 23.0%; Pred. No. 1.5e-56;
Matches 397; Conservative 260; Mismatches 544; Indels 525; Gaps 53;

QY 552 LFTLTAIYSVTLTKAVVREKETRLDTRAMGLSRVLMWLGWFL----- 597
Db 201 LLFAVCMMPVISVARALVVEKSS-VKPYLTITIGLPLMWFLYHLEHFLGVKNTFLTILLS 259
QY 598 -----SC-----LGPFLLSAALLVLVLKGLDILPYS-----HPG 626
Db 260 TLYIFSMDCNCTPYVLGAGIFMVTCHCVSILCTSLPFGKRIVEGMVITLIIAMHLS 319
QY 627 VVFLFLAFAVATVTTQSLLSAFFPSRANLAACGG-----LAYFSLYL 669
Db 320 LEFEDWLFVPLLPNPNYSKLFDATFLASGPNGTPTSALFSKKKTLQSAAVVFGIMI 379
QY 670 --PYVLCA-----WDRRLPAGGRVAASLLSPVAFPGCESLALLE 708
Db 380 SCTVVMVAALFMEXLYTFVGHAFKFRWR-----ILGFSKGKRSKIE 422
QY 709 EGGGAQMHNVGTRPTADVFLSAQVS-----GLLLDAAALYLATWYLEAVCPQOYGP 762
Db 423 ERGDGVE-----DRSTILQCKTEVEGRGSAIADIELSGLVKVY-----QNG-- 463
QY 763 EPWNPFRSRYWCGRPRPKSPAPCTPLDPKVVEAPGLSPGVSVRSLEKRFPGSPOP 822
Db 464 -----BK 465

QY 823 ALRGLSLDFYQGHITAFLGHNGAGKTTTSLSGLFPSPGSAFILGHVDRSSMAIRPH 882
Db 466 AVNGLSLRAIRGOVSILLGHNGCGKSTTFGMITGHWQATEGKVMWIGGDANANRAEAREL 525
QY 883 LGVCPQNVLPDMLTVDEHVMFYGRKLKLS-AAVVGPEQDRLLQDVLGVLSKQSVOTRHLS 941
Db 526 IGYCPQNPYIDELTVWEHLRLVNALKGRSGSDPKMDAESLLKQIELTKRNTLAKNLS 585
QY 942 GGMOKLSVAFAVGGSSQVILDEPTAGVDASRGIIWELLKYEGRTRLIILSTHLDPA 1001
Db 586 GGMKRLCVCNMTGGSRVILLDEPTAGMDPSAKIDVQNMALVKADRTILLTTHYMDEA 645
QY 1002 ELLGDRVAVAGGRLLCCCGSPFLFRRHLGSGVYLTIV--KARLPLTTNEKADTME--- 1055
Db 646 EKLGDWIFVMHSGKWAASGSKHYLKQYGGGMLLTLVFKSVHDPMPRPRKSYETAYDVCKT 705
QY 1056 --GSVDTTREQKNGSQGSRVGT-----PQLLALVQHWVPGARLVBELPHELVLVLYTGA 1108
Db 706 VCSALVKDERGQMTIEISILETKSRSLPTLLKILE-----SVMEE-----DY 747
QY 1109 HDGSFATLFRDLTFLAELRLTGYISDTSLEELPKVVECAADTMDGSCGQHLCGTG 1168
Db 748 NNPEFOALEPDIQEKCRTELEATTIGVSMSSLEQVFIKIGDEC---DDIMNGT----- 796
QY 1169 IAGLD-----VTLRKMPQQTALENGEPAGSAPETDQSGPDVAVRQVGWAL 1216
Db 797 --GVDKKTTEROEKSTLVQYKIQPK-----QGSFK 825
QY 1217 TRQQLQALLKRFLLARRSRRLGFAQIVLPALFVGLALVSLIVPPFGHYPALRLSPTMY 1276
Db 826 LMWVWALLQKRAYVLYENPVQITLIILPLLTLWLFAV-----PFLRLEP--- 871
QY 1277 GAQVSFFSEDPADPGRARLLEALLQEALEPPVQHSRRFSAPEVPAEVAKVLAGNW 1336
Db 872 --KPKLSDIESFDPQYPHSTVLLQLENENDDRLANLNSFSNFVE---VFKTLG----- 922
QY 1337 TPESPSACQCSQPCARRELLPDCPAAAGGPPPPQAVTGSGEVQNLTCGRNLSDFLVKYTP 1396
```

```
Db 923 -----FIVKVNK 929
Qy 1397 R-----LVROGLKTKWVNEVYGGFSLGGRDPGLPSGOELGRSVEELWALLSPLPGA 1450
Db 930 KGDskFYKISQGDKNAAIIMNIIASAMVL--RDPsvTK-----965
Qy 1451 LDRVlKNLTAWSLDAQDSIKIWFNN-----KGWHSVAFVFNPNASNAIIRAHLPQPAR 1505
Db 966 ----LPHVT-----SRVIMNDPKIKYGLASFLLFEN-----994
Qy 1506 HAHsITTLNHNPLNLTKEQLSEALMASSVDVLVSICVWFAMSFVPASFTIVLIEERVTRA 1565
Db 995 -----IFFLLVLAGIFIQS--TVLIEEKICKF 1020
Qy 1566 KHLQlMGGLSPTLYWLNFMCMCNLYLPACIVVLIeLAFQORAYVAPANLPALL--LLL 1623
Db 1021 AHQYLTLGSLTIAYWGVVFLWDF-----LLFTFLLYTIGFLISFVLQGHIEIVV 1072
Qy 1624 LLYG--WSITPLMYPASFFSFVSATVVLTCINLFIGINGSMATFVLELFSDQKLOEVS 1681
Db 1073 IFYGLLFYFAPLVLTLSALINTPTRGNFLL---YMFCCIPWLAYSIVSELHNPFPPIQKYS 1129
Qy 1682 RILKQVFLIF-PHFCLGRGLIDMVR---NQAMADAFERLGRQFQSPLRWEVVGKLLA 1736
Db 1130 DEIEYGRIFNPISIGFLAGLWKIAALNYPKSGLDKHEHETNLWTYEGIFFE-----LM 1183
Qy 1737 MVIQGPLLFLFTLLL-----QHRsQLLPQPRVRSLEPLLGEDEEDVARER 1780
Db 1184 FLFGGIFL--TILGCATLKPFRACFRGTRRSQ--PREKR--YKGIESCkAVKEE 1236
Qy 1781 ERVVOGATQGD-VLVLNLTkVYRGQMPADVRLCLGIPGECFGLLVGNAGKtSTFRM 1839
Db 1237 EQLVQEVDRKNETVLVIDGLVKDF--GKFRAVNDLSISVGHIECFGMILGANGAGKtTTFDI 1294
Qy 1840 VTGDTLASRGEAVLAGHSVAREPSAAHLsMVCYCPQSDAIFELLTGREHLELLARLGVPE 1899
Db 1295 ITGLTMTGGSATIDGHDTET-----IHIGYCPQFDAMIQQISCRQTLRIKALQGYP- 1348
Qy 1900 AQVAGTAGSLARLGLSWYADRPAGTYSGGNKRKLATALAVGDPVAVVFLDEPTTGMDPS 1959
Db 1349 -NVKEVVVELVDCVGMsDFGKLVKNCsGGQKRKISVGIAlMSRATCIILDEPTAGIDPR 1407
Qy 1960 ARRLWNSLLAVNREGR-----SVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKG 2014
Db 1408 ARREIWD-----IHEMEQAKCSIVLTSHSMECEALCTRIGILRKGMIALGTSQSLKS 1463
Qy 2015 RFAAGHTLTLRVPAARS-OPAAAfVAAEFPGSELREAHGG--RLRFOLPFGRCALARV 2070
Db 1464 QYGNTYWMTLLNLSLEDIESVCVIVSEMPDAVLKTPESSLTTSIVWELPKSKSDKWSEK 1523
Qy 2071 FGELAVHGAHBGVDFSVSQTMLEEVFLYFSKQGDKDEDEEQKEA 2116
Db 1524 YNQVEVLAKKANAKDYMLTQASLEDTFIRLI-----TTEEBEEA 1562
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:24:44 ; Search time 27.7979 Seconds
(without alignments)
5119.750 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMYRRR.....QHPKRVSLFDDPSTAETVL 2146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext. 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5768.5	51.8	2261	US-09-526-193A-1	Sequence 1, Appli
2	5747.5	51.6	2261	US-09-032-438C-118	Sequence 118, App
3	4936	44.3	2273	US-09-032-438C-3	Sequence 3, Appli
4	4858	43.6	2235	US-09-032-438C-6	Sequence 6, Appli
5	3818.5	34.3	1375	US-08-665-259-26	Sequence 26, Appl
6	3818.5	34.3	1375	US-08-762-500-26	Sequence 26, Appl
7	2694.5	24.2	1457	US-08-665-259-27	Sequence 27, Appl
8	2694.5	24.2	1457	US-08-762-500-27	Sequence 27, Appl
9	2693.5	24.2	1472	US-09-032-438C-119	Sequence 119, App
10	2507	22.5	1704	US-09-032-438C-120	Sequence 120, App
11	2502	22.5	1704	US-08-762-500-75	Sequence 75, Appl
12	2501.5	22.4	1684	US-08-665-259-25	Sequence 25, Appl
13	2501.5	22.4	1684	US-08-762-500-25	Sequence 25, Appl
14	430	3.9	162	US-10-000-489-102	Sequence 102, App
15	427	3.8	328	US-09-724-797-8	Sequence 8, Appli
16	414	3.7	607	US-09-252-991A-18351	Sequence 18351, A
17	393.5	3.5	588	US-09-489-039A-13579	Sequence 13579, A
18	389	3.5	589	US-09-328-352-7592	Sequence 7592, Ap
19	375	3.4	594	US-09-543-681A-5528	Sequence 5528, Ap
20	374.5	3.4	308	US-09-602-787A-348	Sequence 348, App
21	371.5	3.3	335	US-09-252-991A-20837	Sequence 20837, A
22	363.5	3.3	292	US-09-602-787A-352	Sequence 352, App
23	357.5	3.2	271	US-09-602-787A-350	Sequence 350, App
24	354.5	3.2	315	US-09-328-352-4388	Sequence 4388, Ap
25	350.5	3.1	788	US-09-252-991A-28171	Sequence 28171, A
26	335	3.0	317	US-09-489-039A-10626	Sequence 10626, A
27	333.5	3.0	929	US-09-252-991A-22946	Sequence 22946, A

28	332	3.0	291	4	US-09-107-532A-4205	Sequence 4205, Ap
29	328	2.9	309	4	US-09-252-991A-21204	Sequence 21204, A
30	327	2.9	323	4	US-09-489-039A-12496	Sequence 12496, A
31	319	2.9	922	4	US-09-489-039A-8938	Sequence 8938, Ap
32	316.5	2.8	315	4	US-09-134-000C-6449	Sequence 6449, Ap
33	315.5	2.8	345	4	US-09-252-991A-31957	Sequence 31957, A
34	314.5	2.8	532	4	US-09-543-681A-4646	Sequence 4646, Ap
35	309	2.8	248	4	US-09-710-279-3218	Sequence 3218, Ap
36	306.5	2.8	316	4	US-09-543-681A-6184	Sequence 6184, Ap
37	301	2.7	1280	2	US-08-583-276-19	Sequence 19, Appli
38	300	2.7	1280	4	US-09-767-594-2	Sequence 2, Appli
39	300	2.7	1280	4	US-09-672-810-2	Sequence 2, Appli
40	300	2.7	1280	4	US-09-672-810-5	Sequence 5, Appli
41	300	2.7	1280	4	US-09-672-725C-7	Sequence 7, Appli
42	300	2.7	1280	6	5206352-4	Patent No. 5206352
43	300	2.7	1283	4	US-09-672-810-4	Sequence 4, Appli
44	297	2.7	391	4	US-09-252-991A-20275	Sequence 20275, A
45	297	2.7	1280	2	US-08-752-447-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-526-193A-1
; Sequence 1, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pimstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526,193A
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-526-193A-1

Query Match 51.8%; Score 5768.5; DB 4; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;

Qy	1	MAFWTQLMLLWKNFMYRRQPVLLELLWPLFFILVAVRSHHPLEHCHFPNKP	60
Db	1	MACPQRLLLWKNLTFRRQTQQLLELVAMPFLFILISVRLSPYPPYEQHCHFPNKA	60
Qy	61	LPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRSLNFNDLSVRLADARTVLGGASAH	120
Db	61	MPSAGTLPWVGILCNANNPCFRYPTFGAEGVGVGNFNKSVARLFSDDARLLLSQKDT	120
Qy	121	TLAGLGLIATLRAARSTAQ-----PQTKQSPLEPML--	154
Db	121	SMKDMRKVRLTQQIKKSSNLKQLDFLVNDTEFGFLYHNLSPKSTVDKMLRADVILH	180
Qy	155	-----DVAEL-----LTSLL	164
Db	181	KVFLQGYQLHLTSLCNGSKSEMIQLGDQEVSELCGLPREKLAARVLRNSMIDLKPI	240
Qy	165	RT--ESLGLALGQAQEPHLSLEAAEDLAQELLALRSVLRLR---ALLQRPRTSGP	219

Db 241 RTLNSTGPPFSKELAEATKTLHLSGLTALQELFMRWSMDRQWEMFLTNVNSSSTQI 300
Qy 220 ---LSEALCVRGSPSTVSGPSLWYASDLMLVGV---QEPESALPDSSLSPACSELIG 272
Db 301 YQAVSRIVCGHPEGGLKIKSLWYEDNNYKALFGNGTEEDAETFYDNSTTPYCNDLMK 360
Qy 273 ALDHPISRLRLWRRLKPLILGLKLPADPTTPTRKLMQAQVNRTPPEELTLRLDRVREWMLG 332
Db 361 NLESSPLSRIITKALKPLLVGLKILYTPDTPATROVMAEVNKTFOELAVFHDLEGMBELS 420
Qy 333 PRIETWNSDNVAMQLRLQMQDEGR-RQPRPGRD-----HWEALRSFLDP 379
Db 421 PKIWTFMENSGMDLVRLDSDRNDHFWEQQLDGLDWTAQDIVAFLAKHPEDVQS---- 476
Qy 380 GSGG--XSWODAHADVGLVGTGLRVTECLSLDKLEAAPGEAALVSRAQLALAEHRFWAG 437
Db 477 -SNGSVITWREAFNETQAI RTISRFMECVNLKLEPIATEVFWLINKSMELLDERKEFWAG 535
Qy 438 VVFLGPDSSDPTBHTPDGLGPHVRJIKRMDIDVTRTKIRDRFWDGPAADPLTLR 497
Db 536 IVFTGITPGSIELPH-----HVYKIRMDIDNVERTNKIKDGYWDGPRADPPEDMR 587
Qy 498 YVNGGFVVLQDLVERAAVRVLSGANPRAGLYLOMPYPCYVDVFLVLSRSLPLFLTLA 557
Db 588 YVNGGFAYLQDVVEQAIIRVLITGTEKKTGYTMQMPYPCYVDDIFLVRNGSRMPLFMILA 647
Qy 558 WIYSVTITVAVREKETRLDRMTRAMGLSRAVLWGLFSLCGLPFLLSAALLVLVLKLG 617
Db 648 WIYSVAVIKGI VYEKEARUKETWIRINGLONSILWFSFISLLIPLAVSAGLLVVLKLG 707
Qy 618 DIIPYSHPGVFLPLAFAVATVTOQSPILLSAPFSRANLAAACGLGAYFSLYPLVLCVAV 677
Db 708 NLLPYSDSVVFLSVFVAVVITLQCEFLISTFSRANLAAACGGIITYFTLYPLVLCVAV 767
Qy 678 RDRLPAGGRVAASLLSPVARGCESLALLEEGEGAGMNVGRPT-ADVESLAQVSGL 736
Db 768 QDYVGTILKIPASLLSPVARGCEYFALFEEQIGVQWDLNLFESPVEEDGFNUTTSVM 827
Qy 737 LLDDAALYGLATYLEAVCGQYGIPEPWNPPFRSRYWCGPRPPKSPACPTPLD-PKVL 795
Db 828 MLDEPTLYGVMTWYIEAVFPQYIGIPRPWYFPCTKS YWFGESDEKSHPGSNQKRISEIC 887
Qy 796 VBEAPPLSGVSVRSLEKPPGSPQALRGLSLDFYQCHITAPLGHNGAGKTTLSILS 855
Db 888 MEEPTHLKLGVSIGNLVKYVRDGMKVAVDGLALNFYEGQITSPLGHNGAGKTTTMSILT 947
Qy 856 GLPPPGSGSAFILGHDYRSMMAARPHLGVCPQYVNLFDMLTVDEHVWFYGRKLGLSAAV 915
Db 948 GLFPPTSGTAYILGKDIRSEMSTIRQLGVCPCQHNVLFDMLTVVEHIFWYARLKLSEKH 1007
Qy 916 VGPEQDRLLQDVLG-VSKQSVQTRHLSGGMOKLSVAIAFVGGSQVVLDEPTAGVDVPAS 974
Db 1008 VKAEMEQMALDVLGPPSKLSKTSQLSGGMOKLSVALAFVGGSKVVLDEPTAGVDVPYS 1067
Qy 975 RRGIWELLKYREORTIILSTHLDLAEILGDRVANVAGRLCCCGSPFLRHRHLGSGYY 1034
Db 1068 RRGIWELLKYRQRTIILSTHMDLADVLGDRITAIISHGKLCCVCGSSFLKKNQLGTGY 1127
Qy 1035 LTLVKARLPLTN-----EKADTDMESGVDRQEKNGSQSGRVGTPQLLALV 1082
Db 1128 LTLVKDVESSLSCRRNSSSTVSLKKEDESVSQSSDAGLSGHESDTLIDVSAINLI 1187
Qy 1083 QHWVPGARLVBEELPHELVLVLPYTGADHSGSFATLRELDTRLAELRLTYGISDTSLEEI 1142
Db 1188 RKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRLSDIGISSYGISSETTLEEI 1247
Qy 1143 FLKVVECAADTDMEDGSCCGHLCTGIAGLDVTLRLKMPQETALE--NCEPAGSAPETD 1200
Db 1248 FLKVAEESGVDAETSDGTLFARRNRRAFG-DKQSLCPPTTEDDAADPNDSIDIPESRET 1306
Qy 1201 QGSGPDVAG--RVQGWALTQQLALLKRLFLARRSRGLFAQIVLPALFVGLALVFSL 1258

RESULT 2

US-09-032-438C-118
; Sequence 118, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir

Db 1307 LLSMDGKGSYQVKGWKLTOQQFVALLWKLLIARRSRKGFPAQIVLPAVFCIALVFSL 1366
Qy 1259 IVPPFGHYPALRLSPMTYGAQVFFSDEADPGDPCRARLLLEALLQEAG-----LEE 1308
Db 1367 IVPPFGKYPSELEQPMWYNEQYTFVSNDAPEDTGTLLEALLTKDPGFGTRCMEGNIPD 1456
Qy 1309 PPVQSHSRFSAREVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPPCPAAGGPPP 1368
Db 1427 TPCQAGEEBEETAPVQOTIMDLFQNGNWTWQNFSPACQSSDKIKMLPVCPPGAGLPP 1486
Qy 1369 POAVTSGEVVQNLTCGNLSDFLVKTYPRLVROGLTKTKWVNEVRYGGSISG-GRDPGLP 1427
Db 1487 PQKQNTADTLQDUTGRNISDYILVKTYVQIIAKSLKWKIWNFEFRIYGGFSLGVSNTOALP 1546
Qy 1428 SGQELGRSEBELWALLSPLPGGALDRLVKNLITAWHSLDQDSLSKIWFNNKGHSMVAFV 1487
Db 1547 PSQEVNDIAIKQMKHKLKLANDSSADRFNLNSLGRFMTGLDTRNNVWVFNKNGWHAISFL 1606
Qy 1488 NRASNAILRAHLPPGPARHAHSITTLNHLNLTKEQLSEAAALMASSVDVLVSVICVWAMS 1547
Db 1607 NVINNAILRANLQKGNPSHYGITAFNPLNLTQQLSEVALMTSDVDVLVSVICVIFAMS 1666
Qy 1548 FVPASFVLVLIBERTVTRAKHLQMLMGLSPTLYWLGNFWDMCNLYVPACIWLIFLAFQ 1607
Db 1667 FVPASFVFLIQRVSKAKHLQFISGVKPVYIWLNSFWDMCNLYVPATVILIFICFQ 1726
Qy 1608 RAYVAPANPALLLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATF 1667
Db 1727 KSYVSNLNLVLAALLLYGWSITPLMYPASFFVKIPSTAYVVLTSVNLFIGINGSVATF 1786
Qy 1668 VLELFSQDKLOEVSRIILKQVLIFFPHFCLGRGLIDMVRNOAMADAPERLGRQFQSLRW 1727
Db 1787 VLELFTDNKLNNDILKSVFLIFPHFCLGRGLIDMVRNOAMADALERFGENREVSPLSW 1846
Qy 1728 EVYGNLWAVIOGPIFLFLLTLLQHSOLLPOPRVRSLLPLGEEDESDVARERERVQGA 1787
Db 1847 DLVGRNLFAMAVGVVFLITVLIQVRFPRPRVNAKLSPLNDEDESDVARERQRIIDGG 1906
Qy 1788 TQGDVLVRLNLTKVYRQRMVAVDRLCLGIPGCEGGLLVNGAGKTTSTFRMVTGDTLAS 1847
Db 1907 QONDILEIKELTKIYRRKRKPAVDRI CVGIPGCEGGLLVNGAGKSTTFKMLTGDTTVT 1966
Qy 1848 RGEAVLAGHSVAREPSNAHLSCYQSDAIFELLTGREHLELLARLGRVPEAQVAQTAG 1907
Db 1967 RGDALFNKNSILNIHEVHQNMGYCQFQDAITELLTGREHVEFFALLRGVPEKEVGKVE 2026
Qy 1908 SGLARGLSYADRPAGTYSGGNKRKLATALALVGDPAVVFLEDEPTTGMDPSARRFLWNS 1967
Db 2027 WATKLGVLVYGEKYAGNYSGGNKRKLSTAMALIGPPVVFLEDEPTTGMDPKARRFLWNC 2086
Qy 1968 LLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTURVP 2027
Db 2087 ALSVVKEGRSVLTSHSMECEALCTRMAMVNGRFRCLGSOHLKMRFGDGYTIVVRIA 2146
Qy 2028 AARS--OPAAAFVAAEFPGSELREAHGRLRFPOLPGRCALARVFCELAVHGAHEGVED 2085
Db 2147 GSNPDLKVPQDFGLAPPGSVLKEKRNMLQYL-PSSLSLARIFSIQSCKRHLIED 2205
Qy 2086 FSVSQTMLEEVLYFSGDKQKDE---DTERQKEAGVGDVPAPGLQHPKRVSOFLDDPSTA 2142
Db 2206 YSVSQTLTLDQVFNFAKQSDDDLKXLSLHKQTV-VDVAV-----LTSFLQDEKVK 2257
Qy 2143 ET 2144
Db 2258 ES 2259

APPLICANT: Sun, Hui
APPLICANT: Lupski, James R.
APPLICANT: Nathans, Jeremy
APPLICANT: Anderson, Kent L.
APPLICANT: Leppert, Mark
APPLICANT: Dean, Michael
APPLICANT: Singh, Nanda
APPLICANT: Shroyer, No. 6713300h F.
APPLICANT: Smallwood, Philip M.
APPLICANT: Allikmets, Rando
APPLICANT: Lewis, Richard A.
APPLICANT: Li, Yixin
TITLE OP INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
TITLE OP INVENTION: Transporter And Methods Of Screening For Agents That Modify
TITLE OF INVENTION: ATP-Binding Cassette Transporter
FILE REFERENCE: BYLR-0065
CURRENT APPLICATION NUMBER: US/09/032,438C
PRIORITY FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: US 60/039,388
PRIORITY FILING DATE: 1997-02-27
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.2
SEQ ID NO 118
LENGTH: 2261
TYPE: PR1
ORGANISM: Mouse
US-09-032-438C-118

Query Match 51.6%; Score 5747.5; DB 4; Length 2261;
Best Local Similarity 49.8%; Pred. No. 0;
Matches 1146; Conservative 370; Mismatches 588; Indels 195; Gaps 27;
QY 1 MAFQTQMLLMKMFYRRQVQLLVLLVLPFLFLLVAVRHSHPLEHECHFPNKP 60
DB 1 MACKQLRLMLKWNFTFRRTQCQLLELVAMPFLPLILISVRLSPYPPQHECHFPNKA 60
QY 61 LPSAGTVPMLQGLICNVNNTCPQPTGPEGRGLSNFNDLSVRLADARTVLGASAH 120
DB 61 MPSAGTLPWQGLICNVNNTCPRYPTGPEAGVGVGNFKSIVSRFLSDAQLRLLYSQRT 120
QY 121 TLAGLGLKIATRAAR-----STAQOPTKQSPLEPPM--- 153
DB 121 SIKDMHVLRLMRQIKHPNSNLKQDFLVNDFTFSGFLQHLNLSLPRSTVDSLLQKNVGLQ 180
QY 154 -----LDVAELL----- 160
DB 181 KVFQGYQLHLASLNCNGKLEBIIQGDAEVSALCGLPRKLDAAERVLRYNMDILKPVV 240
QY 161 -----TSLRTESLGLAQOEPLHSLLEAAEDLAQELLALRSVLRLR---ALQRPGR 212
DB 241 TKLNTSHLPTQHLA-----EATVLLDSLGLAQELFTKSWSDMRQEVMLTNVNS 293
QY 213 TSGPLEL---LSEALCSVRGSPSTVGPSLNMYEASDLMELVG-----QBPESALPDSSLS 265
DB 294 SSSSTQIYQVSRIVCGHPGEGGLKIKSLNMYEDNNYKALFGNNTEDVDVTFYDNSTTP 353
QY 266 ACSELIGALDHSPLSLRLWRLLKPLILGKLLPAPDTPETRLMAOVNTEFEELTLRDVR 325
DB 354 YCNDLMKNLESPLRIIWKALKPLVGLKILYTPDTPATROVMAEVNKTFOELAVFHOLE 413
QY 326 EWEMLGPRIFTFMNDSSNVAMQLRLQWQ--DEGRROPRPGRRDHMEALRSFL----- 377
DB 414 GWEELSPOIWFEMNSQEMDLVRLTLLDSRGNDQFWEQKLDGLDWTADQIMAFKAPED 473
QY 378 --DPGSGYSQODAHADVGLVGLTGRVTECLUSDKLEAAPSEAAALVSRAQLLAEHFW 435
DB 474 VQSPNGSVYTWREAFNETNQAIQTISRFMEVCVNLKLEPIETVRLINKSMELDERKFW 533
QY 436 AGVFLGPEDSSDPTEHPTDLG--PGHVRKIRMDIDVTRTKIRDRFWDPCPADPL 493
DB 534 AGIVFTG-----ITPDSVELPHHVKKIRMDIDNVERTKIKDGYWDPGRADPF 583
QY 494 TDLR YVMGGFVYQLDLVERAAVRVLSGANPRAGLYLQOMPYPYVDDVPLRVLRSPLF 553

DB 584 EDMRYVMGGFAYLQDVVEQAIIIRVLVTGSEKKTGVVYQQMPYCYVDDIIFLRVMSRSMPLF 643
QY 554 LTLAWIYSVTLTKAVVREKETRLDDTWRAMGLSRAVILWGLFSLCLGFLISAALLVLV 613
DB 644 MTLAWIYSVAIIKSIIVYKEARLKTWRIMGLDNGILWFSWFSLSIPLLSAGLLVVI 703
QY 614 LKLGILPVSHPGVVFLFLAAFAVATVTSFLLSAFFSRANLAAACGGGLAYLSLYLPPYL 673
DB 704 LKLGILLPYSDPSVVFVFLSVFAMVTILQCFILISTLFSRANLAAACGGIYITLYLPPYL 763
QY 674 CVAMRDLRPAAGRVAASLLSPVAFGFGCESLALLEEQEGEGAGQHNVTGRPT-ADVFSLAQ 732
DB 764 CVAMQDYVGFISIKIFASLLSPVAFGFGCEYFALFEEQIGVQWMDNLFESVPEEDGFNLT 823
QY 733 VSGLLLLDAALYGLATWILEAVCPQYGIPEBWNPFRRSYWCGPRPKSPACPTPLD- 791
DB 824 AVSMMLFTFLYGVNWTYIEAVFPQYGIPIRPWYFPCTKSYWFGBEIDEKSHPSQKGV 883
QY 792 PKVLVEEAPGLSPGVSVRSLEKRPFGSPOPALRGLSLDFYOGHITAFIGHNGAGKTTL 851
DB 884 SEICHEEPTHLRLGVISQNLVKVYRDGMKAVDGLALNFYEGQITSFLGHNGAGKTITM 943
QY 852 SILSGLFPPSGSAFILGHDRVSSMAAIRPHLGVCPQYNVLFDMLTVDHVMFYGRKGL 911
DB 944 SILTGLFPPTSGTAYILGKDIRSEMSIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKL 1003
QY 912 SAAVGPPEQDRLLQDVL-VSKQSVOTRHLSGMORKLSVAIFAVGGSGOVILDEPTAGV 970
DB 1004 SEKHVKAEMEOMALDVGLPSPKLSKTSQSLSGMORKLSVALAFVGGSKVILDEPTAGV 1063
QY 971 DPASRRGIWELLKYREGRTILSTHLDLAEELIGDRVAVVAGGRLCCGSPFLFRRHGL 1030
DB 1064 DYSRRGIWELLKYRGRTILSTHLDLAEELIGDRVAVVAGGRLCCGSPFLFRRHGL 1123
QY 1031 SGYYTLVKARLPLTTN-----EKADTMEGSDVTRQEKNGSQGSRVGTGP 1078
DB 1124 TGYVTLVKDVESSLSCRNSSSTVSCUKKEDSVSQSSSDAGLSGHESDFTLTDVSAI 1183
QY 1079 LALVQHVVPGARLVSELPHELVLVLPYTGANHDGSPATLPRELDTLRLBLRGTGYSIDTS 1138
DB 1184 SNLRKHVSEARLVEDIGHELTYVLPYBAAKEGAPVELFHEIDRSLDSLGSISETT 1243
QY 1139 LEEFLKVVBERCAADTDMEDG-----SCQHLCTGIAGLDVTLRLKMPQ 1183
DB 1244 LEEFLKVAESGVDAETSDGTLPARRNRRAFGDKQSC-LHPFTEDDAVD-----PN 1294
QY 1184 ETALENGEPAGAPETDQSGPDAVG--RVQGWALTROQLQALLKRLFLARRSRGLFA 1241
DB 1295 DSDID---PESRETDLSSGMDGKGSYQLKGWKLTCQQFVALLWKRLIARRSRKGFPA 1349
QY 1242 QIVLPALFVGLALVPSLIVPPFGHYPALRLSPTMYGAQVSPSEDAPODGRARLEALL 1301
DB 1350 QIVLPAPVFCIALVPSLIVPPFGHYPSLELQPMWYNEQYTFVSNADAPDMGTQELNALT 1409
QY 1302 QEAG-----LEEPPVQHSRRFSAPEVAEVAKVLASGNWTPSPSPACCSQPG 1351
DB 1410 KDPGFTRCMGNFIPTPCLAGEEDWTISPVQSIVDLFQNGNMTKNPSPACCSQDK 1469
QY 1352 ARLLPDCPAAAGPPPPQAVTGSGEVVQNLTGRNLSDFLVKTYPRLVKQGLTKKKWNE 1411
DB 1470 IKKMLPVCCPAGGGLPPQKQKTADILQNLTRNIDYLVKTYVQIIAKSLKNKIWNE 1529
QY 1412 VRYGGSGLG-GRDPLGSGQBLGRSVEBELWALLSPGLGALDRVLKNTTAHSHLDAQDS 1470
DB 1530 FRYGSPSLGVNSOALPSEHEVNDIAIKMKLLKLTSDRFLSLSLGRFMAGLDTKN 1589
QY 1471 LKIFWNNKGHSMVAFVNRASNAIILRAHLPPGPARRHAHSITTLNHLNLTKEOLSEALM 1530
DB 1590 VKVFNKNGHMAISSFLNVINNALDRANLQGENPSQYIGITAFNHLPLNTKQQLSEVALM 1649
QY 1531 ASSVDVLVSIQVVFAMSPVPSFTLVLTIEERTVRAKHLQMLMGLSPTLVYLGNFWDNMCN 1590

```
Db 1650 TTSDVVLVSCVIFAMSFVPASFVFLIQRVSKAKHLQFISGVKPVYIWLNSFVWDMCN 1709
Qy 1591 YLVPACIVWLIFLAFOORAYVAPANLPALLILLLLYGWSITPLMYPASFVFFSPSTAYVV 1650
Db 1710 YVVPATLVIIFIGFOQKSVSSITPLVALLLLLLYGWSITPLMYPASFVKIPSTAYVV 1769
Qy 1651 LTCINLFIGINGSMAITVLELFSQKQOEVSRIUKQVFLIPPHCLGRGLIDMVRNOAMA 1710
Db 1770 LTSVNLFIGINGSVATVLELFTNNKLNINDILKSVELFIPPHCLGRGLIDMVRNOAMA 1829
Qy 1711 DAFERLGDROFOSPLRVEVGNLMAVIOGPLELLETLLOHRSQLLPOPRVRSPLLG 1770
Db 1830 DALERFGENRFVSPUSMDLVGRNLFAMAVEGVVFLITVLIQYRFFRPRPVKAKGPLN 1889
Qy 1771 EEDDVARERERVQOGATGDVLVRLNLTKYRGORMPAVDRLCLGIPPGCEGFLGNG 1830
Db 1890 DEDEDVRRQRILDGGQNDLILKELTKYRRKRKPAVDRCIGIPPGCEGFLGNG 1949
Qy 1831 AGKSTFRMTVGTDLARGRAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLEL 1890
Db 1950 AGKSTFRMTVGTDLARGRAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLEL 2009
Qy 1891 LARLGVPKQVQVAGTGLARGLSMVADRPACTYSGNKRKLATLALVGDPAVVFLD 1950
Db 2010 FALLRGVPEKEVGFGGWAIRKGLVYKEIASNYSGNKRKLSTAWALIGBPVVFLD 2069
Qy 1951 EPTTGMDPSARRFLWNSILLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQ 2010
Db 2070 EPTTGMDPKARRFLWNCALSIVKEGRSVLTSHSMECEALCTRWAIWNGRFRCLGSPQ 2129
Qy 2011 HLKGRFAAGHTLTLRVPAARS--OPAAAFAVAEPGSELREAGGRURPOLPGRCAL 2068
Db 2130 HLKGRFAGHTLTLRVPAARS--OPAAAFAVAEPGSELREAGGRURPOLPGRCAL 2188
Qy 2069 RVFELAVHGAEGVEDFVSOTMLEVELYFSKQCKDE--DTEQKEAGVGVDPAPG 2125
Db 2189 RIFSLQSKRKLHIEDYSVQTLDQVFNFAKQDSDHLKDLSLHKNQTV--VDVAV- 2246
Qy 2126 LQHPKRSQFLDDPSTAET 2144
Db 2247 -----LTSFLQDEKVES 2259

RESULT 3
US-09-032-438C-3
; Sequence 3, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattnet, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupeki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2273
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-3
Query Match 44.3%; Score 4936; DB 4; Length 2273;
Best Local Similarity 44.3%; Pred. No. 0;
Matches 1029; Conservative 367; Mismatches 646; Indels 282; Gaps 39;
Qy 1 MAFWTQMLLLWKNFMYRRRQPVQLLVLLWPLFLFFILVAVRHSHPLPHEHCHFPNKP 60
Db 1 MGFVRQIQLLLWKNWTLRKQKIRFVVELVWPLSLFLVLIWLNANPLYSHHECHFPNKA 60
Qy 61 LPSAGTVPMLOGLICNVNNTCFPOLIPGEEPGRLSNFNSLSRLADARTVLLGSAHR 120
Db 61 MPSAGMLPWLQGLFCNVNNTCFQSPPTPGSPGVISVNNYSILARVYRDFOLLENNAPESQ 120
Qy 121 TLAG-----LGLKLATLRA----- 134
Db 121 HLGRITWELHILSQFMDTLRTHPERTAGRIRDIRDLKDEETLTLFLIKNIGLSDSVYL 180
Qy 135 -ARSTAQPOPTKOSPLEPPMLDVA----- 157
Db 181 LINSQVRPEQFAHGVPDPLAKDIACSEALLERFIIFSQRGAKTVRVYALCSLSQGLQWI 240
Qy 158 -----ELTSLRTESGLAQOQEPHLSLEAAEDLAQELLALRSVEL 203
Db 241 EDTLYANVDFKLFRLVPTLLDSRSQGINL-----RSGGGILSDMSPRIQBIHRPSMODL 296
Qy 204 ----RALLQR--PRGTSGLLELSEALCSV-RGPSSTVGPSSLNWNYEASDLMEVLGQBPES 256
Db 297 LWTRPLMQNGGPEPTFKLMGILSDLLCGYPEGGSRV-LSFNWYEDNNYKAFGLIDSTR 355
Qy 257 ALP-----DSSLSPACSELIGALDSHPLSRLMLRRLKPLILGKLLFAPDTPFTTKLMAQVN 312
Db 356 KDIYSYDRRTTSFCNALIQSLESNPLTKIAWAAKPLLMGKILYTPDSPAPARLLKXAN 415
Qy 313 RTEBELTLRLDVRREVEMGLPRIFTFMNDSNVAMLQRLI-----QWQDEGRQ 361
Db 416 STFEELHVRKLVKAMEEVGQIWIYFFDSTQNMIRDTLGNPTVKDFLNRQJGEGIT- 474
Qy 362 PRGGRDHMEALRSFLDPG-----SGOYSQMDAHADVGLVGLRVTVECLSLDKLE 413
Db 475 -----AEAILNFKYKGPRESQADDMANFDRDIFNITDRTLRLVNVQLECLVDKFE 526
Qy 414 AAPSEALVSRALQLLAHRFWAGVFLGPEDSDPTHEHTPDLGPGHVRKIRKMDIDVV 473
Db 527 SYNDETQLTQALSLLEENFWAGVVF-----PDMYPWTSSLPHPVKYKIRMDIDVV 578
Qy 474 TRTNKIRDREMDRCPAADPLTLRLYVWGGFVYLQDLVERAAVRLSGANPRAGLYLOMP 533
Db 579 EKNTKIKORYWDSGPRADPVEDFRYIWWGGPAYIQDMVVEQGITRSQVQAEAPVGIYLOMP 638
Qy 534 YPCVDDVFLRVLRSPLFLTLAWIYSVTLTKVAVVREKETRLRDTMRAMGLSRAVLWL 593
Db 639 YPCFVDDSFMIILNRCFPIFWLAWIYSVMTVKSIVLEKELRKLKTELKNQGVSNVIMC 698
Qy 594 GWFSLCGLPFLLSAALLVLVLKGLDILPYSHGVPVFLFAAFVAVTVTQSFLLSAFSPRA 653
Db 699 TWFLDGSFIMSISIFLLTIIMHGRILHYSDPFLFLFLAFASTATIMLCFLLSSTFFSKA 758
Qy 654 NLAACGGLAYFSILYPLVLCVAVWRDLRPLAGGRVAASSLSPVAFGFCESLALLEEGEG 713
Db 759 SLAAACSGVITYFLYPLPHILCFQAWQRMATLKKAVSLSPVAFGFGTEYLVRFEESGLG 818
Qy 714 AQWNVGTRPT-ADVFSLAQVSGLLILDAALYGLATWYLEAVCPGOYGIPEPNWFFPFRS 772
Db 819 LQMSNIGNSPTGEDFSFLSMQWMLLDAAYGLLAWLQVFPDGYGTPLPWYFLLAQES 878
Qy 773 YWCGPRPKSPAPCPT-----PL-----DPK-----VLVEEAPGSLSPGVSVR 810
Db 879 YWLGGE-----GCSTREERALEKTEPLETETEDPEHGEIHDSFFEREHPGWVPGVCVK 932
Qy 811 SLEKRFPGSPQALRGISLDFYQGHITAFIAGHNGAGKTTTILSILSGLFPSPSGGSAFILGH 870
```


APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IG5-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1375 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-665-259-26

Query Match 34.3%; Score 3818.5; DB 3; Length 1375;
Best Local Similarity 54.1%; Pred. No. 0;
Matches 754; Conservative 224; Mismatches 348; Indels 69; Gaps 14;

Qy 796 VEEAPPGLSPGVSRLEKRPSPQALRLGLSLDPYOGHITAFIGHGAGKTTTSLLS 855
Db 2 MEEBETHURLGVSTONLVKYVRDGMKVAVDGLALNFYEGQITSLFIGHGAGKTTTMSILT 61
Qy 856 GLFPSPGSAFTLGHDRSSMAAIRPHLGVCPOYNVDFMLTYDBHWFYGRLLKLSAAV 915
Db 62 GLFPPTSTAYILGKDIRSESSIRONLVCVPOHNVDFMLTYDBHWFYGRLLKLSKX 121
Qy 916 VGPEDRLQDVLG-VSKQSQVQTHLSGGMQRKLSVAIFVGGSQVILDEPTAGVDPAS 974
Db 122 VKAEMEOMALDGLPPLSKTSQSLSGMQRKLSVALAFVGGSKVILDEPTAGVDPYS 181
Qy 975 RRGWELLKLYREGRTTILSHHDEALLGDRVAVAGRLCCGSPFLRRHLGSGYY 1034
Db 182 RRGWELLKLYRQGRITLSTHMDADILGRIAITSHGLCCVGGSLFLKQLGSGYY 241
Qy 1035 LTLVKARLPLTN-----EKADTMEGSDVDFRQKNGSQSRVGTTPOLLALV 1082
Db 242 LTLVKDVESSLSCRNSSSTVSLCKEDSVSQSSDAGLSDHSDTLTIDVSAI SNLI 301
Qy 1083 QHWVPGARLVELPHELVLVLPYTGADGSGFATLFRDLTRLAELRLTGYSIDTSLEE 1142
Db 302 RKHVEARLVEDIGHLEFVLPYEAKEGAFVELFHEIDRLSDLSGYSIGSETTLEE 361
Qy 1143 PLKVVEECAADTDMEDG-----SCQHLCTGAGLDVTLRLKMPPEAL 1187
Db 362 FLKVAEESGVAETSDGTLPARNRRAFGDKQSC-LHPFTEDDAVD-----PNDSDI 412

Qy 1188 ENGEPAGSAPETDQSGPDVAG--RVQGWALTROOLQALLKXRFLLARRSRGLPAQIVL 1245
Db 413 D-----PESRETDLSSGMDGSGYQLKGWLTQQQFVALLWRLLIARRSRGPPAQIVL 467
Qy 1246 PALFVGLALVFSLIYPPFGHYPALRLSPMTYGAQVSPFSEDAPGPRARLLLEALQEAG 1305
Db 468 PAVFVCIALVFSLIYPPFGKYPSLELQPMWYNEQYTFVSNDAPEQWGTQELLNALTQDPS 527
Qy 1306 -----LEEPPVQHSRFPESAPEVAEVAKVLASGNWTPSPSPACQSCQSGARL 1355
Db 528 FGTRCMEGNPIPDTECLAGEBEDWTISPVPQISVDLFGNGNMTMKQNPSPACQSSDKIKM 587
Qy 1356 LPDCAAGAGGPPPOAVTGSGEVQNLGRLNLSDFLVKTYPRLVRLQGLTKKQWNEVRVG 1415
Db 588 LPVCPGAGGLPPQKOKTADILQNLGRLNLSDFLVKTYVQIIAKSLKNKIWNVEFRVG 647
Qy 1416 GFSLG-GRDPGLPSQELGRSVVEELWALLSPGLGALDRVLKNTLTAHSLDAQDSLKTW 1474
Db 648 GFSLGVSNSQALPPSHEVNDAIKQMKLLKLTQKTSADRFLSSLGGRFAGLDTKNNVVM 707
Qy 1475 FNNKGWHSWAFVNPASNAIILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALWASSV 1534
Db 708 FNNKGWHAISFLNINNAILRANLQNGENPSQYGITAPNHPNLNLTQOQLSEVALMTTSV 767
Qy 1535 DVLVSICVVFAMSFVPASPTLVLIBERVTRAHKLQMGGLSPTLYWLGFLMDMNCNLYVP 1594
Db 768 DVLVSICVVFAMSFVPASFPVFLQERVSXKAGLQPISGVKPVIYWSNFWDMCNVYVP 827
Qy 1595 ACIVVLIFLAFQRAYAPANLPALLLLLLLYGWSITPLMYPASFPFSPVSPYAVVLTIC 1654
Db 828 ATLVIIIFICQOKQSVSSTNLPVALLLLLLLYGWSITPLMYPASFPVFKIPSTAYVVLTSV 887
Qy 1655 NLFINGSMATFVLELPSDOKLQEVSRILKQVELIHPHCLGRLGRLDMVRQNMADAFE 1714
Db 888 NLFINGSVATFVLELFTNNKLANDINDILKSVFLIFPHFCLGRLGRLDMVRQNMADALE 947
Qy 1715 RLGRQPOSPLRWVGNKLLAMVITQGPLLFTLLQHRSQLPQPRVRSPLLGEDE 1774
Db 948 RFGENRFVPSLSDVLGNLFAMAVEGVVFLITVLIQYRFPFIRPVKAKLPPLNDEDE 1007
Qy 1775 DVABERERVQATQGDVLVLRNLTKVYRGORMPAVDRLCLGIPGECFGLLGVNGAGKT 1834
Db 1008 DVRRERQRLDGGQNDILEIKELTKIYRRKRKPAVDRCIGIPGECFGLLGVNGAGKS 1067
Qy 1835 STFRMVTGDTLASGEAVLAGHSVARPSAAHLSMGYCPQSDAIFELLTGREHLLARL 1894
Db 1068 TTFKMLTGDTPTVTRGDAPLKNKNSILSNITHEVQNMGYCPQFDATITELLTGREHVEFFALL 1127
Qy 1895 RGVPDAOVAQTAGSGLARLGLSWADRPAGTYSGNKRKLATALALVGDPAVVFLEDEPT 1954
Db 1128 RGVPKEVKGEGEWAIRKGLVKYGEKYSYSGNKRKLSTAMALIGGPPVVFLEDEPT 1187
Qy 1955 GMDPSARRFLWNSLLAVVREGSVMLTSHSWECEALCSRLAIMVNGRPRCLGSPQHLKG 2014
Db 1188 GMDPKARRFLWNCALSIKVEGRSVMLTSHSWECEALCTRMALIMVNGRPRCLGSPQHLKN 1247
Qy 2015 RFAAGHTLTLRVPAARS--QPAAPFAAEFPSELSREAHGGRRLRQLPBGRCALARYG 2072
Db 1248 RFGDGYTIVRIAGSNPDLPKVPQBFGLAPFGSVLKEKRNMLQYL--PSSLSSLARIFS 1306
Qy 2073 ELAVHGAHGVEDFVSQTMLEEVFLYFSKQDQKDE---DTEEOKEAGVDPAPGLQHP 2129
Db 1307 ILSQSKKRLHIEDYSVQTTLDQVFNFAKQSDDDHLKDLSLHKNQIV--VDVAV----- 1360
Qy 2130 KRVSQFLLDDPSTAE 2144
Db 1361 --LTSFLQDEKVKES 1373

RESULT 6
US-08-762-500-26
; Sequence 26, Application US/08762500
; Patent No. 6030806

; GENERAL INFORMATION:
 ; APPLICANT: Landes, Gregory M.
 ; APPLICANT: Burn, Timothy C.
 ; APPLICANT: Connors, Timothy D.
 ; APPLICANT: Dackowski, William R.
 ; APPLICANT: Van Raay, Terence J.
 ; APPLICANT: Klinger, Katherine W.
 ; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
 ; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
 ; NUMBER OF SEQUENCES: 83
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENZYME CORPORATION
 ; STREET: One Mountain Road
 ; CITY: Framingham
 ; STATE: Massachusetts
 ; COUNTRY: United States of America
 ; ZIP: 01701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/762,500
 ; FILING DATE: 09-DEC-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/665,259
 ; FILING DATE: 17-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/10469
 ; FILING DATE: 17-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dugan, Deborah A.
 ; REGISTRATION NUMBER: 37,315
 ; REFERENCE/DOCKET NUMBER: IGS-9.3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (508) 872-8400
 ; TELEFAX: (508) 872-5415
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1375 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-762-500-26

 Query Match 34.38; Score 3818.5; DB 3; Length 1375;
 Best Local Similarity 54.18; Pred. No. 0;
 Matches 754; Conservative 224; Mismatches 348; Indels 69; Gaps 14;

 QY 796 VEEAPGLSPGVRSLEKFPQSPALRGLSLDFYQGHITAFPLHNGAGKTTTILS 855
 DB 2 MEEPTHLRGLVSLQNLVRYRDMKVAVDGLALNFVEGQITSLFHLNGAGKTTTILS 61

 QY 856 GLFPPSGGSFILGHVRSMAAIRPHLGVCPQNVLFDMLTVDVHEWVFYGRKLGLSAV 915
 DB 62 GLFPPSTGTAYILGKDIRSEMSSIRQLNGVCPQHNVLFDMLTVEEHIWFWYARLKLSEKH 121

 QY 916 VGPEQDRLLODVGIL-VSKQSVQTRHLGGWOKLSVAIAFVGGSQVVLDEPTAGVDPAS 974
 DB 122 VKAIMEQMAIDVGLPPSLKSKTSQSLSGGWOKLSVALAFVGGSKVVLDEPTAGVDPYS 181

 QY 975 RRGWELLKLYREGRTILSLTHLDEALLGDRVAVVAGRLCCCGSPLFLRRHLGSGYY 1034
 DB 182 RRGWELLKLYRQRTILSLTHHDEADILGDRILAIISHGKLCCVGSLSFLKQLGTGY 241

 QY 1035 LTLVKARLPLTTN-----EKADTMEGSDVTREKKNKSGQSGRVGTPQLLALV 1082
 DB 242 LTLVKDVESSLSSCRNSSSTVSLCKEDSVSQSSDAGLGSDESDTLAIIDVSAISNLI 301

 QY 1083 QHWVPGARLVLEELPHELVLVLPYTGADHGSFATLFLRELDLTLABELRLTGTGISTDLEE 1142

DB 302 RXHVSERLVEDIGHELTVLVPYEAKEGAFVELFHEIDRLSLDLGSISSYSETTLEE 361
 QY 1143 FLKVVEECAADTDMEDG-----SCQHLCTGTAGLDVTLRLKMPPOETAL 1187
 DB 362 FLKVABESGVDAETSDGTLFARRNRRAFGDKQSC-LHPFTEDDAVD-----PNDSDI 412

 QY 1188 ENGEPAGSAPETDQSGPDAVG--RVQGNALTRQQLQALLKFLFLIARRSRRLFLAQIVL 1245
 DB 413 D-----PESRETDLLSGMDKGSGYQDKGWKLTOQQFVALLWKLLIARRSRKGFQAIVL 467

 QY 1246 PALFVGLALVFLIPLVPPFGHYPALRLSPWYGAQVSFFSEDAFGDPRARLLLEALLQEAG 1305
 DB 468 PAVFVCIALVFLIPLVPPFGKYPSLELQPMWYNEQYTFVSNDAPEDMGTQELLNALTDPG 527

 QY 1306 -----LEBPPVQHSRPSAPEVPAEVAKVJASGNWTPESPSPACQSQOPARRL 1355
 DB 528 FGTRCMEGNPIPDTPCLAGEEDWTISFPVQSIVDLTFQNGNWTMKNPSPACQSSDKIKKM 587

 QY 1356 LPDCAAGGPPPPQAVTGSGEVVQNLTCGRNLSDFLAKTYPRLVROGLTKTKWNVNRYG 1415
 DB 588 LPVCPGAGGLPPQKQKTADILQNLTCGRNLSDFLAKTYPRLVROGLTKTKWNVNRYG 647

 QY 1416 GFSLG-GRDPGLPSGQELGRSVEELWALLSLPFGGALDRVLKNTLTAHSLDQAQSLKIW 1474
 DB 648 GFSLGVNSQALPPPSHEVNDAIKQMKLLKLTQDTSADRFLSLGRFMAGLDTKNNVKW 707

 QY 1475 FNNKGHSMVAVNVRASNAILRAHLPFGPARIAHSITTLNHLNLTKEQLSEALMASSV 1534
 DB 708 FNNKGWHAISFLNINNAILANLQKGNPSQYGTAFNHPNLNTKQQLSEVALMTTSV 767

 QY 1535 DVLVSLCVVFAMSFPASFTLVLIERTVTRAKHLQMLGGLSPTLYWLNFLWDMCNLYVP 1594
 DB 768 DVLVSLCVVFAMSFPASFPVFLIQRVSKAKHLQFISGVKVIYVLSNFWDMCNLYVP 827

 QY 1595 ACIVWLIFLAQOAYVAPANLPALELALLLXGWSITPLMYPASFVFFSVSTAYVLTCTI 1654
 DB 828 ATLVIIFICFQKSVSVSTNLPVALLLXGWSITPLMYPASFVFKIPSTAYVLTSTV 887

 QY 1655 NLFINGSMATVLELFDQKQEVSRILKQVFLIFPHFCLGRGLIDWVRNQAMADAFE 1714
 DB 888 NLFINGSMATVLELFTNNKLNINDILKSVELIFPHFCLGRGLIDWVRNQAMADALE 947

 QY 1715 RLGDROFQPLRWEVVGKNLLAMVIOGPLFLLTLLQHSQLLPOPVRSLPLGDEDE 1774
 DB 948 RFGENRVSPLSMDLVGRNLFAMAVEGVVFLTLVLIQYRFFRPRPVKAKLPPLNDEDE 1007

 QY 1775 DVARERVVQATQGDVLVRLNLTKVYRGQMPAVDRCLGIPPGECFCGLLGVNGAGKT 1834
 DB 1008 DVREERQILDDGGQNDLLEIKELTKIYRKRKPAVDRCIGIPPGECFCGLLGVNGAGKS 1067

 QY 1835 STFRMVTGDTLASRGAVALAGHSVAREPSAAHLSMGYCPQSDAIFELLTOREHLELRL 1894
 DB 1068 TTFKMLTGDTPTVRGDAFLNKNLSILSNIHEVQNMGYCPQFDAITELLTOREHVEFFALL 1127

 QY 1895 RGYPEAQVAQTAGSLARGLSWYADRPAGTYSGGNKRKLATALALVGDPAVFLDDEPTT 1954
 DB 1128 RGYPEKEVGKFGWARKLGLVKYIGEKYASNYSGNKRKLUSTAWALIGPPVFLDDEPTT 1187

 QY 1955 GMDPSARRFLWNLSLLAVVREGRSVMLTSHSMECEALCSLAIMVNGRFRCLGSPQHLKG 2014
 DB 1188 GMDPKARRFLWNCALSVKGRSVMLTSHSMECEALCTRMAIMVNGRFRCLGSPQHLKN 1247

 QY 2015 RFAAGHTLTLRVPAAAS--QPAAAFVAABEPFGSELREAHGRLRFLQPPGRCALARVFG 2072
 DB 1248 RFGDGYTIVVRIAGSNPDLKPVQEFFGLAPFGSVLKEKHNMLQYQL-PSSLSSLAIRIFS 1306

 QY 2073 ELAVHGAEGHVEDFSYQTMLEEVFLYFSKQCKDE---DTERQKEAGVGVDPAGLQHP 2129
 DB 1307 ILSQSKKRLHIEDYSVSTQTLDOQVFNFAKQDQDDHLKDLSLHKQTV-VDVAV----- 1360
 DB 2130 KRVSQFLDDPSTAET 2144
 : : : : :

Db 1361 --LTSFLODEKVKES 1373

RESULT 7
US-08-665-259-27
; Sequence 27, Application US/08665259
; Patent No. 6028173
; GENERAL INFORMATION:
; APPLICANT: Landee, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1457 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-665-259-27

Query Match 24.2%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 2.4e-231;
Matches 513; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 796 VEEAPPGLSPGVSRSLEKRPFGSPQALRGLSLDFYQGHITAFIUGHNGAGKTTLSILS 855
Db 1 MEEETHPLPVVCDKLTQVKNDRKALNKLNLNENQVVSFLGHNGAGKTTTMSILT 60

Qy 856 GLFPSPGSAFILGHDPVSSAAIRPHLGVCPOYNVLFDMLTVDHVVWFYGRKLKLSAAV 915
Db 61 GLFPPTSGSATYGHDIRTEMDEIRKNLGMCPQHNVLFDRLTVEHLWFYSLKLSMAQEE 120

Qy 916 VGPEODRLQDVLVSQSVQTRHLSGGMKRLSVAIAFVGSGSVVILDEPTAGVDPASR 975
Db 121 IRKETDKMIEDELSNKRHSLVQTLSGGKRLSVAIAFVGSGRAIILDEPTAGVDPASR 180

Qy 976 RGIWELLKLYRGRTLLSTHLLDEAELLGRVAVVAGGRGCCGSPFLRHLGSGYYL 1035
Db 181 RAIWDLILKYPGRTLLSTHLLDEADLLGRDIAIISHKGLKCCGSPFLKGAYXDGURL 240

Qy 1036 TLVKAARLPLTTNEKADTDMESGVDTRQEKNGSQSRVGTQTALLALVOHWTPGARLVEEL 1095
Db 241 TLVKAARLPLTTNEKADTDMESGVDTRQEKNGSQSRVGTQTALLALVOHWTPGARLVEEL 1095

Qy 241 TLVKAARLPLTTNEKADTDMESGVDTRQEKNGSQSRVGTQTALLALVOHWTPGARLVEEL 1095

Qy 1096 PHELVLVLPYTCAGHDGSPATLFRFLDTRIAELRLTGYGISDTSLEIFLKVVEE----- 1149
Db 293 STLSYIILPSEAVKGAERLFQOLEHSLDALHLSFGMLMDTTLBEVFLKVSEEDOSLEN 352

Qy 1150 CAADT-----DMEDGSCQGHILCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETDQSGG 1204
Db 353 SEADVKEKRDVLPAGELTAVGGQAG-NLARCSELAQSASQASSVGS-A-RGEGTG 410

Qy 1205 -----PDVAG-----RVQGWALTQQLQALLKRF 1230

Db 411 YSDGYGYRPLFDNLQDPDNVLSQAEAEALAQVQGSRKLEGMWLMKRFHGLLVKRFH 470

Qy 1231 LARSRRLGFAQIVLPALFVGLALVFSLVIPFGHYPALRLSPMTY-----GAQVSFF 1283

Db 471 CARRNSKALCSQILLPAFFVCVAMTVALSVEIGDPLPLVLSPSQYHNYQTPRGNFIPVA 530

Qy 1284 SED-----APGDPGRARLEALLQAG-----LEEP-----PVQHSN 1316

Db 531 NEERQEVRLRLSP-DASPOQLVSTFRLPSGVGATCVLKSPPANGSLGPMNLNLSGSRLLA 589

Qy 1317 -RESAPEVPAEVAKVLASGNWTPBSPSA----- 1344

Db 590 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVXPDEDSLQAWNMSLPPTAGBETWTS 649

Qy 1345 -----CQCQSGARRLLPDCPAAAGGPPPPQAVTSGSEVVQNLTGRNLSDFL 1391

Db 650 APSLPRLVHEPVRCTCSAQGTGP---SCPSSVGGHPQOMRVV-TGDILTDTGHNVSEYL 705

Qy 1392 VKTVPRLVRQGLTKKWNVEVYGGFSLGGRRDPLPS--GQELGRSVBELWALLSLPBG 1449

Db 706 LFTSDRFLH-----RYGAITFGNVQKSIPIASFARVPPMVRKI----- 744

Qy 1450 ALDRVLKNTAWAHSLDAQDSLKIFWNKNGHSMVAFVNRASNATLRAHLP- --GPARRH 1506

Db 745 AVRRV-----AQVLYNNKGHSMPTYLNLSNNAI LRLANLPKSKNPA-- 786

Qy 1507 AHSITTLNHLNLTKEQLSEALMASSVDVLVSICVVPFAMSFVPASFTVLIEEVRTRAK 1566

Db 787 AXITVTNHPNKTSSASLSLDYLL-QGTDVIAIFIIVAMSFVPASFVFLVAEKSTRAK 845

Qy 1567 HLQLMGSLSPTYLWGNFLWDMCNVLPACIVLIFLAFQORAYVAPANLPALLLLLLLY 1626

Db 846 HLQFVSGCNPIYWLANYVDMNLNLYVPATCCVILFVFDLPAYTSPNTFPAVLSFLLY 905

Qy 1627 GWSITPLMYPASFFFSVPSTAYVVLTCNLNFIGNGSMATFVLELFS--DOKLQEVSRILK 1685

Db 906 GWSITPIMYPASFWEFVPSAYVFLVINLFIGITATVATFLLQLFEHDKDLKVNSYLK 965

Qy 1686 QVFLIFPHFCLGRGLIDMVRNQAMADAPERLGD-RQFQSPLRWEVVGKILLAMVITQGPLF 1744

Db 966 SCFLIFPNYNLGHGLMAYNEYINEYAKIQDPDKMSPFEMDIVTRGLVAMTVGFGV 1025

Qy 1745 LLFTLLQLHRSQLLPQPR---VRSPLLLGEDEEDVARERERVVQAGTQGVVLNRLNLTKV 1801

Db 1026 PFLTIMCOY--NFLRQORLPVSTKPV--EDVDVVASERQVRVLRGDADNDVMKIENLTKV 1081

Qy 1802 YRGO---RMPAVDRCLGI--PPGECGLLVNGAGKSTTFRMVTGDTLASGEAVLAGHS 1857

Db 1082 YKSRKIGRILAVDRCLGVCPVGBCFGLLVNGAGKSTTFKMLTGDESTTGGEAFVNGHS 1141

Qy 1858 VAREPSAAHLSMGYCPOSDAIFELLTGREHLELRLARLGVPEAQVAQTAGSGLARLGLSW 1917

Db 1142 VLKOLLVQOQSLGYCPQDFVPVDELTAHEHLQYLRACIPWKDEAQVVKWALEKLELTK 1201

Qy 1918 YADRPAGTYSNGNRKLTALALVGDPAVFLDEPTTQMDPSARRFLNWSLLAVVREGS 1977

Db 1202 YADRPAGTYSNGNRKLTALIALIGYAFIFLDEPTTQMDPKARRFLNWLILDIKTGRS 1261

Qy 1978 VMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARS--OPAAA 2036

Db 1262 VMLTSHSMECEALCUTRLAIVNGRHLCLGSIQHLKNRFGDGYMTITVKTSQNVKVVVR 1321

Qy 2037 FVAEEFGSELRREAHGG-----RLRFQLPPGRCALARVFGELAVHGAHGVDFSVSQT 2091

Db 1322 FFNRFP-----EAAHQKTPYQVQYL-KSEHISLAQVFSKMQVGVGLGIEDYSVQT 1375
Qy 2092 MLEEVFLYFKDQKQEDTEQKEAGVGDPA-----GLQHPKRVSOFL-----DDPSTA 2142
Db 1376 TLDNVFVNFAK--KQSDNVEQAEPSLPSPLGLLSLLRPRAPTELRALVADEBEDL 1432
Qy 2143 ET 2144
Db 1433 DT 1434

RESULT 8
US-08-762-500-27
; Sequence 27, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landee, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762.500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IG5-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1457 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-762-500-27

Query Match 24.2%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 2.4e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 796 VESAPPGLSGCVSVRSLEKPPGSPQALRGLSLDFYQGHITAPLGHNGAGKTTTILS 855
Db 1 MESEPTLPLVVCVDKLTQVYKNDKLLANKLSLNLYENQVSPPLGHNGAGKTTTILT 60
Qy 856 GLPPPSGSAFILGHGHDVRSMAAIRPHLGVCQVNVLFDMLTVDDEHWFYGRUKLSAAV 915

Db 61 GLFPPTSGSATTIYGHDIRTEMDEIRKNLGMCPOHNVLFDRLTVEEHUWFVSRLSKMAQOE 120
Qy 916 VGPEDRLLQDVGLVSKSQSVQTRHLSGGMOKLSVAIAFVGGSQVWILDEFTAGVDPASR 975
Db 121 IRKETDQWIEDLESNKRHSLSVQTLSGMKRKLVAIAFVGGSRAILIDFPTAGVDYAR 180
Qy 976 RGIWELLKRYREGRTLILSTHHLDEABLLGDRVAVVAGSRLCCGSPFLRRHILGSGYYL 1035
Db 181 RAIWDLTLKYKPGRTILLSTHMDADLLGRIATISHGKLKCCGSPFLKGAIXDGYRL 240
Qy 1036 TLVKARLPLTNEKADTDMEGSVDTROEKNGSGSRVGTTPOLLALVOHVPGARLYEEL 1095
Db 241 TLVKQPAEPTGSOEPPGLASSPSCPL-----SSCSEPPQSVQFIRKHWASSLLVSDT 292
Qy 1096 PHELVLVLPYTGADHGSFATLFRLEDTRLAELRTGYGIGSDTSLSEETFLKVVBE----- 1149
Db 293 STELSYILPSEAVKGAERLFOOLEHSLDALHLSSEGLMDTTLVEEFLKVKSEEDQSLEN 352
Qy 1150 CAADT-----DMEDGSCGHLCTGIAGLDVTLRKMPQETALENGEPAGSAPETDQSG 1204
Db 353 SEADVKEKRDVLPAGEGLTAVGQAG-NLARCSELAQSQASQSSVSGSA-RGEGTG 410
Qy 1205 -----PDVAG-----RVQGWALTROQLALLKREL 1230
Db 411 YSDGYGYRPLFDNLPDNDVSLQEAEMEALAQVQGSRKLEGMWLRKORPHGLLVKRFH 470
Qy 1231 LARRSRRLPAQIVLPALFVGLALVPFSLIVPPGPHYPALRLSPMTY-----GAQVSFF 1283
Db 471 CARRNSKALCSQILLPAFFVCVAMTVALSVPEIGDLPPLVLSPSQYHNYTPQRCNFPYA 530
Qy 1284 SED-----APGDPGRARLLEALLQEAG-----LEEP-----PVQHSSH----- 1316
Db 531 NEERQEVRLRLSP-DASPOQLVSTFRLPSGVGATCVLKS PANGSLGPMNLNSGESLLA 589
Qy 1317 -RESAPEVPAEVAKVLASGNWTPESSPA----- 1344
Db 590 ARPFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVXPDEDSLQAMNMSLPPTGAPETWTS 649
Qy 1345 -----CQCSQPGARLLPDCPAAAGPPPPQAVTQSGEVVQNLTRNLSDFL 1391
Db 650 APSLPLRVHEPVRCTCSAQGTG-----SCPSSVGGHPQMRVV-TGDLITDITGHNVSEYL 705
Qy 1392 VKTYPRLVROGLTKKWNVEVRYGGFSLGDRDGLPS--QBELGRSVBEELWALLSLPLGG 1449
Db 706 LFTSDRFLH-----RYGAITFGNVQKSIAPSGARVPPMVRKI----- 744
Qy 1450 ALDRVLKNLTAWAHSLDAQDSLKIWNKNGHSMVAFVNRASNAILRAHLP-----GPARR 1506
Db 745 AVRRV-----AQVLNNKGYHSMPTYLNNLNNAILRANLPKSKGNPA-- 786
Qy 1507 AHSITTLNHPNLTKEQLSAALMASSVDLVLSICVVFAMSFPVPSFTLVLIBERVTRAK 1566
Db 787 AXITITVNHMPNKTASLSLDYLL-OQTDVVIATFIIIVAMSFVPASFPVFLVAEKSTKAK 845
Qy 1567 HLQIMGLSPTLYWLNFLWDMCNLYLPACIVVLIPLAQQRAYVAPANIPALLLLLLLY 1626
Db 846 HLQFVSGCNPIVYMLANYWDMNLNLYLPATCCVILFVFDLPAYTSPNTPAVLSFLLY 905
Qy 1627 GWSITPLMYPASFPFSPVSTAYVVLCTINLFIGINGSMATFVLELPS-DOKLQEVSRILK 1685
Db 906 GWSITPLMYPASFPFSPVSSAYVFLIVINLFIGITATVATFVLQLFHDHDKLVVNSYLK 965
Qy 1686 QVFLIFPHFCLGRGLIDMVNRQAMADAFERLGB-RQFQSPLRWEVVGKNLLAMVIOGPLF 1744
Db 966 SCFLIFPNYNLGHLMEMAYNEINEYAKIGQDKMKSPFEDIVTRGLVAMTVGEFVG 1025
Qy 1745 LFTLLLOHRSQLLPQPR-----VRSLLPLGEDEEDVARERVVQCATOGDVLVRLNITKV 1801
Db 1026 FFTITMCQY--NFLRQFQRLPVPSTKV--EDDVVASERQVRLRGDADNDMVKIENITKV 1081
Qy 1802 YRQQ--RMPAVDRCLGI--PPGECFGLGNGAGKTTSTPRMVTGDTLGRGEAVLAGHS 1857

Db 1082 YKSRKIGRILAVDRCLGVCVPCGECFGLLVNGAGKISTFKMLTGDSTTGCGEAFVNGHS 1141
Qy 1858 VAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLCVPEAQVAQTAGSGLARGLSW 1917
Db 1142 VUKDILLQVOQSGYCPQFVDPVDELTAAREHLQLYTLRCIPWKDQAQVVKWALEKLELTK 1201
Qy 1918 YADRPAQTYSGGNKRKLATALALVGDPAVVFLEPTTGMDSARRFLMNSLLAVVREGRS 1977
Db 1202 YADKPAQTYSGGNKRKLATAIALIGVPAPIFLDEPTTGMDSARRFLMNSLLAVVREGRS 1261
Qy 1978 VMLTSHSMBEBCALSRALMNGRPRCLGSPQHLKGRFAAGHTLTLRVPAARS-OPAAA 2036
Db 1262 VMLTSHSMBEBCALSRALMNGRPRCLGSPQHLKGRFAAGHTLTLRVPAARS-OPAAA 1321
Qy 2037 FVAAPFPGSELNEAAGG-----RLRFQPPGRCALARVFGELAVHGAHGVDFSVSQT 2091
Db 1322 FNRNFP-----EAAAGKTPYKQYQL--KSHISLAQVFSKWEQVGVGLIEDYVSQT 1375
Qy 2092 MLEEVLYFSKDGKDETEQKEAGVGVDPAP-----GLQHPKRVSQFL-----DDPSTA 2142
Db 1376 TLDNVFVNAK---KQSDNVEQEAEPSSPLSPGLGILLSLRPRPAPTELRAVADPEDL 1432
Qy 2143 ET 2144
Db 1433 DT 1434

RESULT 9
US-09-032-438C-119
; Sequence 119, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 1472
; TYPE: PRT
; ORGANISM: Mouse
US-09-032-438C-119

Query Match 24.2%; Score 2693.5; DB 4; Length 1472;
Best Local Similarity 40.8%; Pred. No. 3e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

Qy 796 VEEAPPGISPGVSVASLEKRFPGSPQALRGLSLDFYQGHITAFIGHNAGKTTLSILS 855
Db 16 MEEEPHTLPLVVCVDKTKVYKNDKKLAKLSLNLXENQVVSFLGHNGAGKTTTMSILT 75
Qy 856 GLFPSPGSAFILGHVDVSSMAAIRPHLGVCQYQVNLFDMLTVDHVVWFYGLKGLSAV 915
Db 76 GLFPPTSGSATIYGHDIRTEMDEIRKKNLGMCPQHNVLPDRLTVEBHLWFYSRLKSMQEE 135

Qy 916 VGPEODRLLQDVLVSKOSVQTRHLSSGGMORKLSVAIAFVGGSOVVILDEPTAGVDPAQR 975
Db 136 IRKETDKMIEDLELSKNKSHLVQTLSSGGMKRLSVAIAFVGGSRAILDEPTAGVDPAQR 195
Qy 976 RGIBELLKRYREGRTLILSTHHLDBAELLGDRVAVVAGRLCCCGSPFLRRHLGSGYYL 1035
Db 196 RAINDLILKYKPGRTILLSTHMDADLLGDRIATISHGKLKCCGSPFLKGAAYKDGRL 255
Qy 1036 TLVAKARLPLTNEKADTDMEGSVDTQBEKKNGSQSRVGTQQLLALVQHWVPGARLVEEL 1095
Db 256 TLVKQAPABSTQBEGLASSPSCPL-----SSCSEPOVQOFIRKXVASSLLVSDT 307
Qy 1096 PHELVLVLPYTGHAHGPSPATLRELDTRLAELRLTGYGISTDSLSSEIFLKVVEE----- 1149
Db 308 STELSYIILPSEAVKKGAFERLFOQLEHSLDALHLSFGMLDITLSEVFLKVEEEDOSLEN 367
Qy 1150 CAADT-----DMEDGSCGHILCTGAGLDVTLRLKMPPOETALENGEPAGSPAPETDQSGG 1204
Db 368 SEADVKEKRDVLPAGEGLTAVGGQAG-NLARCSELAQSAQSLQASSVGS-A-RGEGTG 425
Qy 1205 -----PDVAG-----RVQGWALTQQLQALLKRL 1230
Db 426 YSDGYDYRPLPDNLQDPDNVLSQAEWELAAQVGGSRKLEGMWLKMRQFHLGLVKRRH 485
Qy 1231 LARRRRGLFAQIVLPALFVGLALVFSLIVPPFGHPALRLSPMTY-----GAQVSFF 1283
Db 486 CARRNSKALCSQILLPAPFVCMVVALSVPEIGDLPPLVLSFSQYHNYTQPRGNFIPYA 545
Qy 1284 SED-----APGDPGRARLLEALLQAG-----LEEP-----PVOHSSH----- 1316
Db 546 NEERQERYRLSP-DASPOQLVSTPRLPSPGVGATCVLKSPPANGSLGPMNLNLSGSRLLA 604
Qy 1317 -RFSAPVPAEVAKVLASGNWTPESSPA----- 1344
Db 605 ARFFDMCLESFTQGLPLSNFVPPPPSPAPSDSPVKPDEDSLQAWNMSLPPTAGPETWTS 664
Qy 1345 -----COCQPGARRLLPDCPAAAGPPPOAVTGSGEVQNLTRNLSDFL 1391
Db 665 APSLRLVHEPVRCTCSAQTGF-----SCPSSVGGHPQWRVV-TGDILDTIGHNVSEYL 720
Qy 1392 VKTYPRVROGLTKKXWNVRYGFSIGGRDPGLPS--GOELGRSVELWALLSPLPGG 1449
Db 721 LFTSDRFLH-----RYGAIITFCNVQKSIIPASFGARVPPMWAKI----- 759
Qy 1450 ALDRVLKNTAWAHSLDQDSLKIFWNNKGWHSVAFVNRASNAIILRAHLPP---GPABH 1506
Db 760 AVRRV-----AQVLYNNKGHSMPTYLNSLNAIILRAWLKPSKGNPA-- 801
Qy 1507 AHSITTLNHLTKQEALMASSVDVLVSIQVFPAMSFVPASFTLVLEIEERVTRAK 1566
Db 802 AYKITVTHNPNKTSASLSLDYL--QGTVDVIAIFIIVAMSFVPASFVVFLVAEKSTAK 860
Qy 1567 HLQMGGLSPTLYMGNFLMDKCNVLPACIVLIFLAPQQGAYVAPANLPAALLLLLY 1626
Db 861 HLQFVSGCNPIVYMLANYVMDMLNLYVPATCCVILFVFDLPAYTSPNTPFPAVLSLELY 920
Qy 1627 GWSITPLMYPASFEFSPSTAYVVLTCINLFTIGNSMATFVLELPS-DOKLQESRLK 1685
Db 921 GWSITPIMYPASFPFEPSSAYFVILNLFITATVATFLLQLFHEHDKLUKVNSYLUK 980
Qy 1686 QVLFIFPHFCLGRGLIDMVRNQAMADAFERLGD-RQFQSPLRWEVVGKNLLAMVQGPFL 1744
Db 981 SCFLIFPNYNGHGLMWAYNEYINEYAKIQQPKMKSPEWMDIVTRGLVAMTVGGFVG 1040
Qy 1745 LLFTLLQHRQLLPQR-----VRSPLLGEDEEDVARERVRVQAGTGDVLVLRNLTKV 1801
Db 1041 PFLTIMCOY--NFLRQORLPVSTKPV--EDVDVASERQVRVLRGDADNDMWKIENLTKV 1096
Qy 1802 YRGO---RMPAVDRCLGI--PPGCFGLGVNGAGKTTSTERMVTGDTLASRGEAVLAGHS 1857
Db 1097 YKSRKIGRILAVDRCLGVCVPCGECFGLLVNGAGKTTSTFKMLTGDSTTGCGEAFVNGHS 1156
Qy 1858 VAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLCVPEAQVAQTAGSGLARGLSW 1917

Db 1157 VLKDLLOVQSLGCPDFVVDLTAHQLYTRURCIPWQDAQVVKWALEKELTK 1216
Qy 1918 YADRPACTYGGNKKRLATLALVGDPAVFLDPTTGMDSARRFLWNSLLAVRGRS 1977
Db 1217 YADKPACTYGGNKKRLSTALIGYPAFLDPTTGMDSARRFLWNSLLDILKTGRS 1276
Qy 1978 VMLTSHNMBECCALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTURVPAARS-QPAAA 2036
Db 1277 VMLTSHNMBECCALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTURVPAARS-QPAAA 1336
Qy 2037 FVAAPFSGSELREAHGG-----RLRFQPLPGRCALARVCELAHGAHGVDFVSQT 2091
Db 1337 FNNRNP-----EHAQKTPYKVOYQL-KSEHISLAQVFSKMBQVVGVLGIEDYSVSQT 1390
Qy 2092 MLEBVLYFSKQDKQDTEBQKAGVDPAP-----GLQPKRVQSQFL-----DPSSTA 2142
Db 1391 TLDNVFNFAK---KQSDNVEQQAEPSSLPGLLSLLRPRPAPTELRAALVADEPEDL 1447
Qy 2143 ET 2144
Db 1448 DT 1449

RESULT 10
US-09-032-438C-120
; Sequence 120, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; FILE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 120
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-120

Query Match 22.5%; Score 2507; DB 4; Length 1704;
Best Local Similarity 33.7%; Pred. No. 2e-214;
Matches 639; Conservative 314; Mismatches 630; Indels 314; Gaps 47;
Qy 307 LMAQVNRTEBELTLARDVREWMLGPRIFTFMNDSSNVAQLRLQWQDGRQRPGG 366
Db 9 LLLKNVNTLQRRKVLVTLELFL---PLLF-----SGLIWLRLKIQSENVNATIYPG- 59
Qy 367 RDHEALRSFLDPGSGYSWQ-----DAHADVHLVGLTGRVTECLSLDKLEAAESEAALV 422
Db 60 -QSIQELPLFTFPFGDGTWELAYIPSHSDAAKVTETVRALVINM-RVRGFPSEKDFE 117
Qy 423 SRAIQLLAEHFWAGVFLPGSDSDTEHPTDLPGLGHVRIKIRMDIDVVTR----- 475
Db 118 DYIRYDNCSSSVLAAVFEHFNHS---KEPLPLAVKXHLRFSYTRNMYMTQTGSFLK 174

Qy 476 -----TNKIRDFWDPGPAADPLT-----DLRYWGVGFVYLQDLVERAAVRLSGAN--- 522
Db 175 ETBGMHTTSLFPLFPNPGP-RETPSPGEGPGYIREGFLAVQHAVDRAIMHEYHADAATRQ 233
Qy 523 --PRAGLYLQOMPYPCVDDVFLRVLRSRLPLFLTLAWIYSVTLTKVAVVREKETRLRDT 580
Db 234 LFORLVTITKRFYPPIEDPFLVAIQYQLPLLLLSFTYTLTATIAVAVVQEKERRLKEY 293
Qy 581 MRAMGLSRVWLWGLFSLCGLPFLLSAALLVLVLKLG-----DILPYSHPGVVFLFLAFA 635
Db 294 MRMMGLSSMLHWSAMFLFLFLFLIAASFMTLLFCVKVKNVAVLSRSDPSLVLAFLCF 353
Qy 636 AVATVTSQSLLSAFRRANLAACGLAYPSLYPLVPLCVAMDRDLRDPAGGRVAASLLSPV 695
Db 354 AISTISFSFMSVSTFFSKANMAAAGGFLYFTYPIYFVAPRYNMWMTLSQKLSCLLSNV 413
Qy 696 AFGCGESLALLEEQEGEAQWNVGTRPTA-DVFSLAQVSGLLLLDAALYGLATWYLEAV 754
Db 414 AMANGAQLIGKFEAKGMIQWRDLSLVNVDDDFCFQVGLGMLLDSVLGLVTWWEAV 473
Qy 755 CPQYGIPEPWNFFRRSYWCGPRPKSPAPCTPLDPKLV-----EEAPPGLSPGVS 809
Db 474 FPGQGVQPQWYFFIMPSYWG-KPRAVAGKEEDSDPEKALRNEYFEASPEDLVAGIKI 532
Qy 810 RSLEK--RFPGPSQPALRGLSLDFYOGHITAFLGHNGAGKTTTSLISGLFPSPGGSFAFI 867
Db 533 KHLKIVFRVGNKDRAAVRDLNLNLYBGOITVLVGHNGAGKTTTSLMLTGLFPPTSGRAYI 592
Qy 868 LGHDIRSSMAAIRPHLGVCPOYNVLFDMLTVDHVMVFYGRKLKGLSAAVVGPEQDRLLQDV 927
Db 593 SGVEISQDMVQIRKSLGCLCPQHDILFDNLTVAEHLIFYAQLKGLSRQKCEEVKQMLHII 652
Qy 928 GLVSKOSVQTRHLSGGMQRKLSVAIAFVGGSQVVLDEPTAGVDPASRRGIWELLKYRE 987
Db 653 GLEDKWNRSRFLSGGMRRKLSIGIALIAGSKVILDEPTSGMDAISRAIWDLLQOKS 712
Qy 988 GRTLILSTHLDEAELLGDRVAVVAGGRLLCCCSPLFLRHLGSGYVLTIVKARLPITN 1047
Db 713 DRTIVLTTHFMDADLLGDRIAIMAKGELCCSSSLFKOKYGAGYHMTLVKE----- 765
Qy 1048 EKADTMEGSDVTRQEKNGSQSRVGTPOLLALVQHWVPGARLVEELPHELVLVLYTGT 1107
Db 766 -----PHCNPEISQLVHHVFNATLESSAGAELSFILPRES 802
Qy 1108 AHGDSFATLRFELDLRLAEURLTYGISTSLBEIFLKV----- 1146
Db 803 TH--RPEGLFAKLEKKQKELGIAFGASITTMBEVFLRVGKLVDSMDIOAIQIPALQYQ 860
Qy 1147 VECAADTMEDSCGQHLCTGIAGLDVTLKMPQETALENGEPAGSAPETDQSGPD 1206
Db 861 HERRASDWAVDNLG-----AMDPSDGTGALIEER-----T 893
Qy 1207 AVGRVQGWALTROLOQALLKRLFLARRSRRLFAQIVLPALFVGLAVFSLIVPPFGHY 1266
Db 894 AVKLTGLALHCOQFWAMFLKKAAYSREWKNVAAQVILPVTCTVLTALLAINYSSELFD 953
Qy 1267 PALRLSPTMYGAQVSFFSDAPGDGPRARLLEALLQEALEEPVQVSHSRFSAPEVPAE 1326
Db 954 PMLRLTLGEXGRTVVPFSV---PGTSQLGQQL-----SEH----- 985
Qy 1327 VAKVLASGNWTPSPSPACOCQSGARRLLPDCEAAGGPPPPQAVTGSVEVQNLGRN 1386
Db 986 -----LKDALQAEQ--OEPREVLG-----D 1003
Qy 1387 LSDPLVKTYPRLVQRGLKTKKWNVEVRYGGFSGLGRDPLGSLGQELGRSVEELWALLSPL 1446
Db 1004 LEFLI-----FRASVEGGGFN----- 1020
Qy 1447 PGGALDRVLKXLTAWAHSLLDAQSLKTIWFNNKGHSMVAFVNRAASNAILRAHLPPGPARH 1506
Db 1021 -----ERCL--VAASFRDVGERTVVAALFNNQAYHSPATALA VVDNLLFK--LLCGP--H 1069

Db 654 LEDXNSRSFSLGMRKLSIGIALIAGSKVLLIDPTSGMDAISPRAIWDLLQKQSD 713
Qy 989 RTILSTHLLDEAELGDRVAVVAGRLCCGSPFLRRHLGSGYLLTVKARLPTTNE 1048
Db 714 RTIVLTTHFDEADLLGDRIAMAKGBLQCCGSSLFLKYGAGYHMTLKE----- 765
Qy 1049 KAUDMEGSDVTQEKNGSGSRVGTPTQLLALVOHWVPCARLVEELPHELVLVLPYGA 1108
Db 766 -----PHCNPEDISQLVHHHPNATLESSAGAELSFILPREST 803
Qy 1109 HDGSFATLFRDLTRIAELRLTGYGSDTSLEBIFLKV-----V 1147
Db 804 H--RFEGLFAKLEKKXELGASFGASITTMEEVFLVGLVDSSMDIQAIQALPQYQH 861
Qy 1148 EECADTDMDGSCGQHLCTGIAGLVDTLRLKMPPOBETALENGEPAGSAPETDQGGSDA 1207
Db 862 ERRASDWAVDNLCG-----AMDPDSGIGALIEBER-----TA 894
Qy 1208 VGRVOGWALTRQLOALLKRLARRSRGLFAQIVLPALFVGLALVFLSLIIPPFGHYP 1267
Db 895 VKLNTGLALHCQOPWAMFLKKAAYSREWKAQAQVLPVLTCTVTLALLAINYSSELDPP 954
Qy 1268 ALRLSPMYGAQVSFFEDAPGPRARILLEALLQEALEBEPVQSHSRFSAPVPAEV 1327
Db 955 MLRLTIGEYRTVVPFSV-----PGTSQLCQL-----SEH----- 985
Qy 1328 AKVLASGNWTPESPSPACQSPQARLLPDCPAAAGPPPPQAVTSGEVWQNLGRLN 1387
Db 986 -----LKDALQAEQ--QEPREVLG-----DL 1004
Qy 1388 SDFLVKTYPLRVROGLTKWVNEVRYGFSGLGRDPGLPSGQELGRSVEELWALLSPLP 1447
Db 1005 EEFLLI-----FRASVEGGGFN----- 1020
Qy 1448 GGALDRVLKNTLAHWSLDQDLSLKIWFNKGWHSVAFVNRASNATLRAHLPPGPARHA 1507
Db 1021 -----ERCL--VAASPRVDGRTVVNFALFNQAHS PATALAVDNLFLK--LCCGP--HA 1070
Qy 1508 HSITTLNHP-----LNLTKQLGSEALMASSVDVLVSVFVAMSFPASPFTVLVIBERV 1562
Db 1071 -SIVVSFPQPSALQAAKQFNEG--RKGFIDIALNL--LFAMAFLASTFSLAVSERA 1124
Qy 1563 TRAKHLQMLGSLTLYWLNFLWDMCNLYVPACIVVLIIPLAFOORAYVAPANIPALLLL 1622
Db 1125 VQAKHVQFVSGVHVASFVLSALLWDLISFLIPSLLLWVFKAFDVFRAFTDRGHMADTLL 1184
Qy 1623 LLLYGWSITPLMYPASFPFSVPSTAYVVLITCINLFIGINGSMAFVLEL--FSDOKLOE 1679
Db 1185 LLLYGWAIIPLMYLMNFFFLGAATAYTRLITIFNLISGI-----ATFLMVTIMRIPAVKLE 1240
Qy 1680 VSRILKQVLIFFPHFCLGRGLIDMVRN-----QAMADAFERLGRDQFQSP-LRWEV 1729
Db 1241 LSKTLDHVLVLPNHLGMAVSSPYENYETRYCTSEVAHAHYCKYNIQYQENFYAWSA 1300
Qy 1730 --VGNLLAMVIOG-----PLFLFTLLLOHRSOLLPPQVR-----SLPLLGEDE 1774
Db 1301 PGVGRFVMAASACAYLILLFLIETNLLQRLGILCALRRRTLTLYTRMPVL-PEDQ 1359
Qy 1775 DVARERERVOGATOG--DVLVLRNLTKVYRGORMP--AVDRLCGIPGCEFCGLLGVN 1829
Db 1360 DVADERTRILAPSDSLHPLIILIKELSKYE--QRVPLLAVDRLSLAVQKCEFCGLLGFN 1418
Qy 1830 GAGKTSTFRMVTGDTLASRGEAVLAGHSVAREPSAHLMSGYCQSDAIFELITGREHLE 1889
Db 1419 GAGKTTTFKMLTGESLTSGDAFVGGRHISDVGKVRQIRGYCQPDALLDHTMGREMLV 1478
Qy 1890 LLARLRGVEPAQVAQTAGSLARGLSWYADRPACTYSGGNKRKLATALALVGDPAVVEL 1949
Db 1479 MYARLRGIPERHIGACVENTLRGLLEPHANKLVRTYSSGGNKRKLSTGIALIGEPVIFL 1538
Qy 1950 DEPTTGMDPSARRFLWNSLLAVVREGRSVMLTSHSMEECALCSRLAIMVNGRRCILGSP 2009
Db 1539 DEPTSGMDPVARRLLMDTVAPARES GKAIITSHSMEECALCTRLAIMVQGGFKCLGSP 1598

Qy 2010 QHLKGFPAAGHTLTTLVPAARSOPA-----AAFVAAEFFPGSELREAHGGRRLRFLPFGGRC 2065
Db 1599 QHLKSGFGYSRLAKVQSEGQOEALBEPKAFVDLTTPGSLVEDEHQGMVHYHL-PRDL 1657
Qy 2066 ALARVGEIAGHGAHGVDFSVQTMLEEVFLYFS 2101
Db 1658 SWAKVGEILEKAKEKYGVDYVSQISLEQVFLSPA 1693

RESULT 12

US-08-665-259-25
; Sequence 25, Application US/08665259
; Patent No.: 6028173
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IG5-9.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-259-25

Query Match 22.4%; Score 2501.5; DB 3; Length 1684;
Best Local Similarity 33.8%; Pred. No. 6.1e-214; Indels 309; Gaps 46;
Matches 635; Conservative 304; Mismatches 628;

Qy 327 VBWMLGPRITFWNDSSNVAMQLQLQMDQEGRRQPRPGGRDHMEALRSFLDPGSGYSW 386
Db 6 VLELFPLPLF-----SGILWLRLKIQSENVPNTIYFG--QSIQELPLFTFFPPGDTW 58
Qy 387 Q-----DAHADVGHVLGTGRVTECLSLDKLEAPSEALVSRALQLLAEHRFAGVWVFLG 442
Db 59 ELAYIPSHSDAAKAVTETVRALVINM-RVRGPFSEKDFEDYTRYDNCSVLAAYVFEH 117
Qy 443 PEDSSDPTHEPTDLGPHVRIRKIMDIDVVTETNKI---RDRFWD-----PGPAD 491
Db 118 PFNHS---KEPLFLAVKYLHLRFSYTRRNVMWTGTGFFLKETGWHHTSLPLFPNPGPR 174
Qy 492 PLT-----DLRYVVGGFVYLQDLVERAAVRLSCAN-----PRAGLYLQOMPYPCYVDDV 541

Db 175 ELTSPDGPGEYIREGLFVAVHQAHDRAIMEYHADAAATRLQFLQRLVTVTIKRPPYPPFIADP 234
Qy 542 FLRVLSRLPLTLTAWIYSVTLTKAVVREKETRLDTRMAMLSRAVLWGLFSLCLG 601
Db 235 FLVAIOYQLPULLLSFYTTALTIRAVVQEKERLKYMMGMGLSSWLHWSANFLFLFL 294
Qy 602 PFLSAALLVLVLKLG-----DILPYSHGVVFLFLAFAVATVTSFLLSAPFSRANLA 656
Db 295 FLIIAASFWTLFCVKVKNVAVLSRSDPSLVLAFLCFAISTISFSFMWSTFFSKANNA 354
Qy 657 AACGGLAYFSLYLPVLCVAMRDRLPAGGRVAAASLLSPVAGFGCESLALBEOGGAOW 716
Db 355 AAFGFLYFFTYIPFFVAPRYNMWTLQKLCCLSNVAMGAQLIGKFEAKGMGIOW 414
Qy 717 HNVGTRPTA-DVFSLAQYSGILLDDAALYGLATWYLEAVCPQYGIPEPMNPPFRSVMC 775
Db 415 RDLSPVNVDDFCGQVGLMGLLOSVLVGLVTWMEAVFPQGVQPVWTFWTFMPSWYC 474
Qy 776 GPRPKSPAPCPTPLDPKVLV-----EBAPGLSPGVSVRSLEK--RFPSPQPALRGLS 828
Db 475 G-KPRAVAGKEEEDSPEKALRNEYFEAPEDLVAGIKIKHLSKVFRVGNKDRAAVRDLN 533
Qy 829 LDFYQCHTAFLHNGAGKTTTSLSGLPFPSSGSAFILHGDVRSMAAATRPHLGVCPQ 888
Db 534 LNLXBSQITVLGHNGAGKTTTSLMTGLFPPTSGRAYISGYEISQDMVQIRKSLGLCPQ 593
Qy 889 YNVLFDMLTVDHVMWFYGRKLKGLSAAVGVPODRLLQDVLGVSKOSVOTRHLSCGMORKL 948
Db 594 HDILFDNLTVABHLFYAQLKGLSKQKCPPEVKQMLHIIIGLEDKNSRSLSGMRRKL 653
Qy 949 SVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKRYEGRTLILSTHHLDEAELLGDRV 1008
Db 654 SIGIALIAGSKVLLDEPTSGMDAISRAIWDLLQROKSDRTIVLTTHFMEADLGDRI 713
Qy 1009 AVVAGRLCCCGSLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKNGS 1068
Db 714 AIMAKGELQCCGSSLFLKQYKAGYHMTLVKE-----745
Qy 1069 QGSRVGTPLQALLAQHVHPGARLVEELPHELVLVLPYTGADHGSFATLRFELDTLAE LR 1128
Db 746 --PHCNPDISQLVHHVHPNATLESSAGALSIFLPRESTH--RPEGLFAKLEKKQKREL 801
Qy 1129 LTGYSIDTSLEEIFLKV-----VEECAADTDMEDGSCGQHLC 1167
Db 802 IASFGASITMEVEFLRVGLVDSSMDIQAQLPALQYQHERRASDVAWDSNLG-----856
Qy 1168 GIAGLDVTLRLKMPQETALENGEPAGAPETDQSGPDVAGRVQGWALTQQOQALLK 1227
Db 857 -----AMDPDSDIGALIBER-----TAVKLTNTGLALHCQCFWAMFLK 894
Qy 1228 RFLIARRSRGLFAQIVLPALFVGLALVSLIVPPFGHYPALRLSPMTYGAQVSPFSEDA 1287
Db 895 KAAYSWRKWKMAQVLPVLCVTLLALAIYNSSELPDDPMLRTLGEYGRVTPVPSV--952
Qy 1288 PGDGRARLLBALQEALEBPPVQHSRRFSAPVPAEVAKVLAAGNWTPESSPAQC 1347
Db 953 ---PQTSQGLQQL-----SEH-----965
Qy 1348 SQPGARRLLPCPAAGPPPPQAVTSGSEVVQNLTGRNLSDFLVKTYPRLVROGLTKK 1407
Db 966 -----LKDALQAG--QEPREVLG-----DLEFLI-----989
Qy 1408 WNEVR YGFSLGGEDPGLPSQOELGRSVBELWALLSPLPGCALDRVLKNTAWAHSIDA 1467
Db 990 FRASVEGGFN-----ERCL--VAASFRDVG 1014
Qy 1468 QDSLKIFNNKGMHSMVAVNRASNAIRLAHLPGPARHAHSITTLNHP-----INLTKE 1522
Db 1015 RTVVNALFNNQAYHSPATALAVVDNLFK--LLCGP--HA-SIVVSNFPQPRSAQAOKD 1069
Qy 1523 QLSAALMASSVDVLVSTCVVPFAMSVFPASFTLVLIBERVTRAKHLQMLGGLSPTLWLG 1582
Db 1070 QFNEG---RKGFDIALNL--LFAMAFLASTFSILAVSRAVQAKHVQFVSGVHVASFWSL 1124

Qy 1583 NFLWDMCNLYVPACIVVLIFLAFOORAYVAPANLPALLLLLLLYGWSITPLMYPASFPFS 1642
Db 1125 ALLWDLISFLIPSLLLLVVFKAFDVRATRGHMADTLLLLLLLYGWAIIPLMYLWNPFL 1184
Qy 1643 VPSTAYVVLTCINIFIGINGSMATFVLEL-----FSDQKLQEVSRILKQVLPFPHFCLGRG 1699
Db 1185 GAATAYTRLTIFNLISGI-----ATFLMVTIMRIPAVKLEELSKTLDHVFVLPNHLGMA 1240
Qy 1700 LIDVRN-----QAMADAFERLGDROFQSP-L-RWEV--VGKNLLAMVIOG-----P 1742
Db 1241 VSSPYENTRYRTSSEVAHYCKYNIQENFYAWSPGVGRFVMAASGCAYLIL 1300
Qy 1743 LFLFTLLQLLQHSOLLQOPVR-----SLPLLGEDEVDVARERERVQATQG---D 1791
Db 1301 LFLIETNLLQRLGILCALRRRTLTETYTMPVL--PEDQDVADERTRILAPSPDSLHT 1359
Qy 1792 VLVRNLTQVYRGQMP--AVDRLCLGIPPOGECFGLLVNGAGKTSTFRWVTGDTLASRG 1849
Db 1360 PLIIKELSKVYE--QRVPLLAVDRLSLAVQKGEFCFGLLVNGAGKTTTFKMLTGESLTS 1418
Qy 1850 EAVLAGHSVAREPAAHLSMGCYCPQSDAIFELLTGREHLELLARLGVPEAQVQATGSG 1909
Db 1419 DAFVGGHRISSDVGVKQRIGYCPQFDALLDHMTGREMLVMYARLGRGIPERHIGACVENT 1478
Qy 1910 LARLGLSWYADRPAGTYSGGNKRKLATALAVGDPVAVVFLDEPTTGMDPSARRFLWNSLL 1969
Db 1479 LRGLLLEPHANKLVITYSGGNKRKLSTGIALIGEPANIFLDEPSTGMDPVARRLLWDIVA 1538
Qy 1970 AVVREGSRVMLTSHMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPA 2029
Db 1539 RARESGKAIITSHMEECEALCTRLAIMVQCFKLGSPQHLKSKFGSGYSLRAKVQSE 1598
Qy 2030 RSQPA---AAFVAABEPGSELREAHGRLRFPQLPGRCALARVFGELAVHGAHEGVED 2085
Db 1599 GQBEALBEFKAFVDLTTPGVSYLEDEHQGMVHYHL--PGRDLSSWAKVFGILEKAKYGVDD 1657
Qy 2086 FVSQTMLEEVFLYFS 2101
Db 1658 YSVSQISLEQVFLSFA 1673

RESULT 13

US-08-762-500-25
; Sequence 25, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIORITY NUMBER: 08/000,000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-762-500-25

Query Match 22.4%; Score 2501.5; DB 3; Length 1684;
Best Local Similarity 33.8%; Pred. No. 6.1e-214;
Matches 635; Conservative 304; Mismatches 628; Indels 309; Gaps 46;

327 VWEMLGPRITTFWNDSSNVAMQLRQLQMQDEGRQRPPGGRDHMEALRSFLDPCSGGYSW 386
6 VLELFLPLLF-----SGILWLRKLTQSENVNATIYPG--QSIQLPLPFTFFPPGDTW 58

387 Q-----DAHADVGHVLTGLGRVTECLSLDKLEAAPSAAALVSRAQLLAAHRFWAGVVF 442
59 ELAYIPSHSDAAKAVTETVRALVINN-RVGRPPSEKDFEDYTRYNCSSSVLAAVVFEH 117
443 PEDSDPTHTPDLPQGHVRIKIRMDIDVVTTNKI---RDRFWD-----PGPAAD 491
118 PFNHS---KEPLAVKYHLRFSYTRNNYMTQTGSFFLKETEGWHTTSLFLFPNPGPR 174

492 PLT-----DLRYVWGFFVYLQDIVERAAVRLSGAN-----PRAGLYLOQMYPYCYDDV 541
175 ELTSPDGGEFGYIREGFLAVQHAVDRAIMEYHADAATRQLFQRLTVTIKRPYPPTIADP 234

542 FLAVLSRSLPLFTLAWIYSVTLTKAVVPEKETRLDTRANGCLSPAVLWLGWFLSCLG 601
235 FLVAIQVQLPLLLLGFTYALTATRAVVOQEKRRUKYRMNMGLSWLHWSAWFLFLFL 294

602 PFLLSAALLVLVLKLG----DILPSHPGWFLFLAAPAVATVTSFLLSAFFSRANLA 656
295 FLIIASFWTLTLCVKVKPNVAVLSRSDPSLVLAFLLCFAISITSFWMVSTFSKANMA 354

657 AACGGIAYFSLYLPYVLCVAWDRDLPAGGRVAASLISPVAFGFCESLALLESQGEAQW 716
355 AAFGGFLYPTYIYPFVAPRYNNMTLSQKLCSSLSNVAMAMGAQLIGKFEAKGMGIOW 414

717 HNVGTRPTA-DVFSLAQVSGLLILDAAALYGLATWYLEANCPCQGYGPEPNWFFRISYWC 775
415 RDLLSPVNVDDDFCFQGVLMGLLSDSLYGLVWTYMEAVPGQVPGPQPFYFIMPSYWC 474

776 GRPPKSPAPCPTLPDPKVLV----EAPPGLSPGVSVLEK--RFGPSPOPALRGLS 828
475 G-KPRAVAGKEEDSDPEKALRNEYFAEPEDLVAGIKIKHLKVFVGNKDRAAVRDLN 533

829 LDPYQGHITAFLGHNGAGKTTTSLISGLPFPSPGGSAPILGHDVRSMAAIRPHLGYCPQ 888
534 LNLVEGQITVLLGHNGAGKTTTSLTGLFPPTSGRAYISGYEISQDMVQIRKSLGLCPQ 593

889 YNVLFDMLTVDVHHVFGRIKGLSAAVVGPEQDRLLQDVLGSKQSVQTHLSGGHQWKL 948
594 HDILFDNLTAEBHLFYFAQLKGLSRQCKPEVKQMLHIIGLEDKWNSRSRFLSGGMRKL 653

949 SVAIAFVGGSQVVLDEPTAGVDPASRRGIWELLLLKYREGRTTILSTHHLDEAELLGDRV 1008
654 SIGIALTAGSKVILDEPTSGMAISRRAIWDLLQROKSDRTVLTITHFHDEADLLGDRI 713

1009 AVVAGGRLLCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSDVTRQEKNGS 1068

```


Db 1599 GQCEALEBFKAFVDLTFPGSVLEDEHQGMVHYHL-PGRDLSWAKVFGILEKAKERYGVDD 1657
QY 2086 FSVSOTMLEEVFLYFS 2101
Db 1658 YSVSQTSLSEQVFLSFA 1673

RESULT 14
US-10-000-489-102
; Sequence 102, Application US/10000489
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US6 DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 102
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-489-102

Query Match 3.9%; Score 430; DB 4; Length 162;
Best Local Similarity 54.7%; Pred. No. 4.3e-30;
Matches 76; Conservative 26; Mismatches 37; Indels 0; Gaps 0;
QY 1 MAFWTQLMLLKNFMYRRQRPVQLLVLLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MACWPQLRLLWKNLTFRRQTCQLLELVANPLFIFLILSVRLSYPPYEQHECHFPNKA 60
QY 61 LPSAGTVPWLOGLICNVNNTCFPQPTGPEEGRLSNFDLSVRLADARTVLGGASAH 120
Db 61 MPSAGTLPWVGIIICNANNPCFRYPTGPEAGVGNGFNKSIVARLFSRRLLLYSQKDT 120
QY 121 TLAGLGKLIATLRAARSTA 139
Db 121 SKMOKRKVLRLTLOQIKKSS 139

RESULT 15
US-09-724-797-8
; Sequence 8, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 328

; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-8
Query Match 3.8%; Score 427; DB 4; Length 328;
Best Local Similarity 34.9%; Pred. No. 2.9e-29;
Matches 112; Conservative 59; Mismatches 142; Indels 8; Gaps 5;
QY 1796 RNLTKVYRGQMPAVDRLCIGIPGCECGLLGVNGAGKTSFRMVVTGDTLASRGEAVLAG 1855
Db 12 RDLVKVFGQTR--AVDGLDLVVRAGTIHGVLGPNAGAKTTAKMLATLMRPTSGTASVLG 69
QY 1856 HSVAREPSAAHLSMGYCPQSDAIFELLTGRHLELLARLGVPEAQVAQTAGSGLARLGL 1915
Db 70 HDVVREAAAEVRRRIGLTGQTMVSDEDMTGVLQNLILAGRLQGLRHASAAAEQMLEAFDL 129
QY 1916 SWYADRPAGTYSGGNKRLATALVGDPAVVFLDEPTTGMDPSARRFLWNSLLAVVREG 1975
Db 130 TEVGGRLVKTFSGGQRRRIDVAASMVVTPELLFLDEPTTGDLPRSRSEWEMIRALVRDG 189
QY 1976 RSVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAG-HTLTLRVPAARSQPA 2034
Db 190 GTVLTTQYLDLADHLADELTLIDHGRIVAGQTPPELKASRAAGVLDVRLRDPERRADAG 249
QY 2035 AAFVAABFPFGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAHGEVDFSVSQTMLE 2094
Db 250 ALLAKAVGAAADL--DSDPARLSVRVTDPDRAALA--LGELARAGIH--VDDFTLQGPSLD 304
QY 2095 EVFLYFSKQDQKDEDTBEQKE 2115
Db 305 TVFLALTGHSTVDASEEEAE 325

Search completed: December 29, 2004, 22:53:49
Job time : 43.7979 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 243.611 Seconds
(without alignments)
5068.549 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQMLLLWKKNFYRR.....QHPKRVQFLDDPSTAEVTVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_spot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	11132	99.9	2146	2 Q9BZC4	Q9BZC4 homo sapien
2	11130	99.9	2146	2 Q81ZY2	Q81ZY2 homo sapien
3	11128	99.9	2146	2 Q9NR73	Q9NR73 homo sapien
4	10249	92.0	2008	2 Q9ES58	Q9ES58 homo sapien
5	8652	77.6	2170	2 Q7TNJ2	Q7TNJ2 rattus norv
6	8600.5	77.2	2159	2 Q8UVV4	Q8UVV4 mus musculu
7	5817	52.2	2260	2 Q8UVV4	Q8UVV4 gallus gall
8	5761.5	51.7	2261	1 ABC1_MOUSE	P41233 mus musculu
9	5751.5	51.6	2261	1 ABC1_HUMAN	Q91V24 mus musculu
10	5510.5	49.5	2201	2 Q80ZB2	O95477 homo sapien
11	5030	45.1	2269	2 Q6T941	Q80ZB2 rattus norv
12	5030	45.1	2269	2 AAR87835	Q6T941 canis fami
13	5029	45.1	2269	2 Q6T940	Aar87835 canis fam
14	5029	45.1	2269	2 AAR87836	Q6T940 canis fami
15	5028	45.1	2269	2 Q6T942	Aar87836 canis fam
16	5028	45.1	2269	2 AAR87834	Q6T942 canis fami
17	4971.5	44.6	2281	2 O02698	Aar87834 canis fam
18	4948.5	44.4	2310	1 ABCR_MOUSE	O02698 bos taurus
19	4948.5	44.4	2310	2 AAR57853	A35600 mus musculu
20	4936	44.3	2273	1 ABCR_HUMAN	O35600 mus musculu
21	3629.5	32.6	2434	1 ABC2_RAT	P78363 homo sapien
22	3621	32.5	2436	2 Q9HC28	Q9HC28 rattus norv
23	3611	32.4	2436	1 ABC2_HUMAN	Q9HC28 homo sapien
24	3496.5	31.4	2434	1 ABC2_MOUSE	Q9BZC7 homo sapien
25	3437	30.8	1487	2 Q8BPY1	P41234 mus musculu
26	3351.5	30.1	1771	2 Q76MW7	Q8BPY1 mus musculu
27	3351.5	30.1	1771	2 BAA83014	Q8BPY1 mus musculu
28	2808	25.2	2595	1 ABI2_HUMAN	Baa83014 homo sapien
29	2715	24.4	1416	2 Q6ZP24	Q6ZP24 homo sapien
30	2715	24.4	1416	2 BAC98084	Q6ZP24 mus musculu
31	2714	24.4	2127	2 Q86W12	Bac98084 mus muscu
					Q86W12 homo sapien

32	2701.5	24.2	5058	2 Q86UQ4	Q86UQ4 homo sapien
33	2645.5	23.7	2143	2 Q80T20	Q80T20 mus musculu
34	2503	22.5	1704	2 BAB86781	Bab86781 homo sapi
35	2502	22.5	1704	1 ABC3_HUMAN	Q99758 homo sapien
36	2471.5	22.2	1538	2 Q80XT2	Q80XT2 mus musculu
37	2468	22.1	1764	2 Q8MUA3	Q8MUA3 strongyloce
38	2463.5	22.1	1704	1 ABC3_MOUSE	Q8R420 mus musculu
39	2316.5	20.8	1563	2 Q7FZY9	Q7FZY9 anopheles g
40	2272.5	20.4	1713	2 Q9VRG4	Q9VRG4 drosophila
41	2184	19.6	1679	2 Q7Q7K8	Q7Q7K8 anopheles g
42	2165.5	19.4	1683	2 Q6XBG3	Q6XBG3 mus musculu
43	2165.5	19.4	1683	2 AAF73044	Aap73044 mus muscu
44	2091.5	18.8	1750	2 Q7QJG5	Q7QJG5 anopheles g
45	2070	18.6	1578	2 Q7Q5K8	Q7Q5K8 anopheles g

ALIGNMENTS

RESULT 1

Q9BZC4	PRELIMINARY;	PRT; 2146 AA.
AC	Q9BZC4;	
DT	01-JUN-2001 (Tremblrel. 17, Created)	
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)	
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)	
DE	ABC transporter member 7.	
CN	Name=ABCA7;	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI TaxID=9606;	
RN	[1] _SEQUENCE FROM N.A.	
RP	MEDLINE=21328888; PubMed=11435699;	
RX	Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,	
RA	Shulenin S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,	
RA	Mattei M.G., Dean M., Denefle P., Chimini G.;	
RT	"Comparative analysis of the promoter structure and genomic	
RT	organization of the human and mouse ABCA7 gene encoding a novel ABCA	
RT	transporter.";	
RL	Cytogenet. Cell Genet. 92:264-270(2001).	
CC	-!- SIMILARITY: Belongs to the ABC transporter family.	
DR	EMBL; AF328787; AAK00959.1; .	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.	
DR	GO; GO:0000166; F:nucleotide binding; IEA.	
DR	GO; GO:0006810; P:transport; IEA.	
DR	InterPro; IPR003439; AAA ATPase.	
DR	Pfam; PF00005; ABC_tran; 2.	
DR	ProDom; PD000006; ABC_transporter; 2.	
DR	SMART; SM00382; AAA; 2.	
DR	PROSITE; PS00211; ABC_TRANSPORTER_1; 1.	
DR	PROSITE; PS00893; ABC_TRANSPORTER_2; 2.	
KW	ATP-binding.	
SQ	SEQUENCE 2146 AA; 234305 MW; 2391728D5AD97E75 CRC64;	
Query Match	99.9%; Score 11132; DB 2; Length 2146;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 2144; Conservative	1; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MAFWTQMLLLWKKNFYRRQVQLLVLLPFLFFILVAVRHSHPLEHHECHFPNKP 60	
DB	1 MAFWTQMLLLWKKNFYRRQVQLLVLLPFLFFILVAVRHSHPLEHHECHFPNKP 60	
QY	61 LPSAGTVPWLOGLICNVNNTCPQPTGCEPGRLSNFNDSLSRLADARTVLGGASAH 120	
DB	61 LPSAGTVPWLOGLICNVNNTCPQPTGCEPGRLSNFNDSLSRLADARTVLGGASAH 120	
QY	121 TLIALGKLIATRAARSTAQAQPTKQSPLEPPMLDVAELTSLRTESLGLALGQAQEP 180	
DB	121 TLIALGKLIATRAARSTAQAQPTKQSPLEPPMLDVAELTSLRTESLGLALGQAQEP 180	

Db 121 TLAGLGLIATLRAASTAQOPTKQSPLEPPMLDVAELLTSLIRTESLGLALQAEPL 180
Qy 181 HSLEAAEDLAQELLALRSIVELRALLQRPRTGSGPLELSEALCSVRGSSSTVGPSSLNW 240
Db 181 HSLEAAEDLAQELLALRSIVELRALLQRPRTGSGPLELSEALCSVRGSSSTVGPSSLNW 240
Qy 241 YEASDLMELVQEPESALPOSSLSPACSELIGALDHPRLSRLWRRLLKPLILGKLLFAPD 300
Db 241 YEASDLMELVQEPESALPOSSLSPACSELIGALDHPRLSRLWRRLLKPLILGKLLFAPD 300
Qy 301 TPFTKRLMAQVNRTEBELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLQMQDEGRR 360
Db 301 TPFTKRLMAQVNRTEBELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLQMQDEGRR 360
Qy 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGTIGRVTECLSLDKLEAAPSEAA 420
Db 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVGTIGRVTECLSLDKLEAAPSEAA 420
Qy 421 LVSRALQLLABHRFWAGVVFILGPDSDPTBHTPTDLGPGHVRIRKIRMDIDVTRTKIR 480
Db 421 LVSRALQLLABHRFWAGVVFILGPDSDPTBHTPTDLGPGHVRIRKIRMDIDVTRTKIR 480
Qy 481 DRFWDPGFAADPLTDLRYVMGFFYLQDLVERAAVRVLSGANPRAGLYLQMPYPICYDD 540
Db 481 DRFWDPGFAADPLTDLRYVMGFFYLQDLVERAAVRVLSGANPRAGLYLQMPYPICYDD 540
Qy 541 VFLRVLSRSLPLFTLAWIYSVTITVKAUVREKSTRIDNRAMGLSRAVLWLGWFLSCL 600
Db 541 VFLRVLSRSLPLFTLAWIYSVTITVKAUVREKSTRIDNRAMGLSRAVLWLGWFLSCL 600
Qy 601 GPFLLSALILVLKIGDILPYSPHGVPVFLPAAFAVATVQSFLLSAFFSRANLAAACG 660
Db 601 GPFLLSALILVLKIGDILPYSPHGVPVFLPAAFAVATVQSFLLSAFFSRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVAVNRDLRACGRVAASLLSPVAFGFCESLALLEEQEGQWENVG 720
Db 661 GLAYFSLYLPVLCVAVNRDLRACGRVAASLLSPVAFGFCESLALLEEQEGQWENVG 720
Qy 721 TRPTADVLSLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPWNPFRRSYWCGRPP 780
Db 721 TRPTADVLSLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPWNPFRRSYWCGRPP 780
Qy 781 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPOPALRGSLDPYQGHITAF 840
Db 781 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPOPALRGSLDPYQGHITAF 840
Qy 841 GHNGAGTKTTLSILSGLFPSPGSAFILGHVDVRSMAAIRPHLVCQYQNVLFDMLTVD 900
Db 841 GHNGAGTKTTLSILSGLFPSPGSAFILGHVDVRSMAAIRPHLVCQYQNVLFDMLTVD 900
Qy 901 HWFYGRBKLSAAVQPEQDRLLQDVLVSKQSVQTRHLSGCMQKLSVAIAFVGGSQV 960
Db 901 HWFYGRBKLSAAVQPEQDRLLQDVLVSKQSVQTRHLSGCMQKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLAEALGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLAEALGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGYLLTLVKARLPLTTNEKADTMEGSVDTROEKNKSGQSRVGTPOLLA 1080
Db 1021 SPLFLRRHLGSGYLLTLVKARLPLTTNEKADTMEGSVDTROEKNKSGQSRVGTPOLLA 1080
Qy 1081 LVQHWVPGARVIEELPHELVLVLPYTGADHSGSFATLFRDLTRLAELRLTGYISDTSLE 1140
Db 1081 LVQHWVPGARVIEELPHELVLVLPYTGADHSGSFATLFRDLTRLAELRLTGYISDTSLE 1140
Qy 1141 EIFLKVVEECAADTMEDSGCQHLCGIAGLDVTLRKWPQPTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTMEDSGCQHLCGIAGLDVTLRKWPQPTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRQGWALTRQIQALLKRLFLIARRSRRLGFAQIVLPALFVGLALVFSLIV 1260
Db 1201 QGSGPDVAVRQGWALTRQIQALLKRLFLIARRSRRLGFAQIVLPALFVGLALVFSLIV 1260

Qy 1261 PPRGHYPALRLSPMTYGAQVSFFSEADAPGDPGRARLLLEALLQEAAGLEPPVQHSHSRFSA 1320
Db 1261 PPRGHYPALRLSPMTYGAQVSFFSEADAPGDPGRARLLLEALLQEAAGLEPPVQHSHSRFSA 1320
Qy 1321 PEVPAEYAKVLASGNWTPESPSPACQSQPCARLLLPDCPAAAGPPPPQAVTSGGEVQ 1380
Db 1321 PEVPAEYAKVLASGNWTPESPSPACQSQPCARLLLPDCPAAAGPPPPQAVTSGGEVQ 1380
Qy 1381 NLTCRNLSDFLVKTYPRLVQGLTKKVVNEVRYGGFSLGGRDPLGSGQELGRSVBELW 1440
Db 1381 NLTCRNLSDFLVKTYPRLVQGLTKKVVNEVRYGGFSLGGRDPLGSGQELGRSVBELW 1440
Qy 1441 ALLSPLPGGALDRVLKNTAWAHSLODQSLKTFWNNKGHSMVAFVNRASNAITLRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSLODQSLKTFWNNKGHSMVAFVNRASNAITLRAHLP 1500
Qy 1501 PGPARHAHSITTLNHNPLNLTKEQLSEAAALMASSVDVLVSCVVFAMSFVPASFTVLIEE 1560
Db 1501 PGPARHAHSITTLNHNPLNLTKEQLSEAAALMASSVDVLVSCVVFAMSFVPASFTVLIEE 1560
Qy 1561 RVTRAKHLQMLGSLPFLYWLGNFLMDMCMYLVPACIVVLI FLAFOQRAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMLGSLPFLYWLGNFLMDMCMYLVPACIVVLI FLAFOQRAYVAPANLPALL 1620
Qy 1621 LLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKQEV 1680
Db 1621 LLLLLYGWSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKQEV 1680
Qy 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDROFQSPLRWEVVKNNLLAMVIQ 1740
Db 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDROFQSPLRWEVVKNNLLAMVIQ 1740
Qy 1741 GPFLFLPTLLQHRSQLLPQVRSLPPLGEEDEDAVARERERVVQCATQGDVLVLRNLTK 1800
Db 1741 GPFLFLPTLLQHRSQLLPQVRSLPPLGEEDEDAVARERERVVQCATQGDVLVLRNLTK 1800
Qy 1801 VYRGQRPVADRCLGIPPCGECFLGVNGAGTKSTFRMTVGTDLASRGEAVLAGHSVAR 1860
Db 1801 VYRGQRPVADRCLGIPPCGECFLGVNGAGTKSTFRMTVGTDLASRGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSNYAD 1920
Db 1861 EPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSNYAD 1920
Qy 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDPSARRFLWNLSLLAVVREGRSVML 1980
Db 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDPSARRFLWNLSLLAVVREGRSVML 1980
Qy 1981 TSHSMEECEALCSPLAIVNNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQPAAAFVAA 2040
Db 1981 TSHSMEECEALCSPLAIVNNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQPAAAFVAA 2040
Qy 2041 EFPGSELREAHGRLRFLQPLPGRCALARYFVELAVHGAHBGVDEDFSVQTMLEEVFLYF 2100
Db 2041 EFPGSELREAHGRLRFLQPLPGRCALARYFVELAVHGAHBGVDEDFSVQTMLEEVFLYF 2100
Qy 2101 SKQQKDEDETEQKEAGVVDPAFLQHPKRVSQFLDDPSTAEVTL 2146
Db 2101 SKQQKDEDETEQKEAGVVDPAFLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 2

Q81ZY2 PRELIMINARY; PRT; 2146 AA.
ID Q81ZY2
AC Q81ZY2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter ABCA7.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=20549028; PubMed=11095984;
RA Kamineki W.E., Piehler A., Schmitz G.;
RT "Genomic organization of the human cholesterol-responsive ABC
RT transporter ABCA7: tandem linkage with the minor histocompatibility
RT antigen HA-1 gene.";
RL Biochem. Biophys. Res. Commun. 278:782-789(2000).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF3111060; AAN04657.1; JOINED.
DR EMBL; AF3111061; AAN04657.1; JOINED.
DR EMBL; AF3111062; AAN04657.1; JOINED.
DR EMBL; AF3111063; AAN04657.1; JOINED.
DR EMBL; AF3111064; AAN04657.1; JOINED.
DR EMBL; AF3111065; AAN04657.1; JOINED.
DR EMBL; AF3111066; AAN04657.1; JOINED.
DR EMBL; AF3111067; AAN04657.1; JOINED.
DR EMBL; AF3111068; AAN04657.1; JOINED.
DR EMBL; AF3111069; AAN04657.1; JOINED.
DR EMBL; AF3111070; AAN04657.1; JOINED.
DR EMBL; AF3111071; AAN04657.1; JOINED.
DR EMBL; AF3111072; AAN04657.1; JOINED.
DR EMBL; AF3111073; AAN04657.1; JOINED.
DR EMBL; AF3111074; AAN04657.1; JOINED.
DR EMBL; AF3111075; AAN04657.1; JOINED.
DR EMBL; AF3111076; AAN04657.1; JOINED.
DR EMBL; AF3111077; AAN04657.1; JOINED.
DR EMBL; AF3111078; AAN04657.1; JOINED.
DR EMBL; AF3111079; AAN04657.1; JOINED.
DR EMBL; AF3111080; AAN04657.1; JOINED.
DR EMBL; AF3111081; AAN04657.1; JOINED.
DR EMBL; AF3111082; AAN04657.1; JOINED.
DR EMBL; AF3111083; AAN04657.1; JOINED.
DR EMBL; AF3111084; AAN04657.1; JOINED.
DR EMBL; AF3111085; AAN04657.1; JOINED.
DR EMBL; AF3111086; AAN04657.1; JOINED.
DR EMBL; AF3111087; AAN04657.1; JOINED.
DR EMBL; AF3111088; AAN04657.1; JOINED.
DR EMBL; AF3111089; AAN04657.1; JOINED.
DR EMBL; AF3111090; AAN04657.1; JOINED.
DR EMBL; AF3111091; AAN04657.1; JOINED.
DR EMBL; AF3111092; AAN04657.1; JOINED.
DR EMBL; AF3111093; AAN04657.1; JOINED.
DR EMBL; AF3111094; AAN04657.1; JOINED.
DR EMBL; AF3111095; AAN04657.1; JOINED.
DR EMBL; AF3111096; AAN04657.1; JOINED.
DR EMBL; AF3111097; AAN04657.1; JOINED.
DR EMBL; AF3111098; AAN04657.1; JOINED.
DR EMBL; AF3111099; AAN04657.1; JOINED.
DR EMBL; AF3111100; AAN04657.1; JOINED.
DR EMBL; AF3111101; AAN04657.1; JOINED.
DR Genew; HGNC:37; ABCA7.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00005; ABC_transporter.
DR ProDom; PD000006; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234421 MM; 33A128082D7B5BAF CRC64;
Query Match 99.9%; Score 11130; DB 2; Length 2146;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2143; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAFWTQMLLLWKNFMYRRQVQLLVLLVLPFLFFILVAVRHSHPPLEHHECHFPNKP 60
DB 1 MAFWTQMLLLWKNFMYRRQVQLLVLLVLPFLFFILVAVRHSHPPLEHHECHFPNKP 60
QY 61 LPSAGTVMWLOGLICNVNNTCPQPTPGSEPRQLSNFNDSLSVRLLDARTVLGGASAH 120
DB 61 LPSAGTVMWLOGLICNVNNTCPQPTPGSEPRQLSNFNDSLSVRLLDARTVLGGASAH 120
QY 121 TLAGLGKLIATLRAAARSTAQPOPTKQSPLEPPMLDVAELLTSLLRTEISGLAQOEPL 180
DB 121 TLAGLGKLIATLRAAARSTAQPOPTKQSPLEPPMLDVAELLTSLLRTEISGLAQOEPL 180
QY 181 HSLLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELLSEALCSVRGSSVTGPSLNM 240
DB 181 HSLLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELLSEALCSVRGSSVTGPSLNM 240
QY 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRLKPLIIGKLLFAPD 300
DB 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRLKPLIIGKLLFAPD 300
QY 301 TPPTKLMQVNRTPPELTLLRDVREVMELGPRIFTFMNDSSNVAMQLQRLQMDEGRR 360
DB 301 TPPTKLMQVNRTPPELTLLRDVREVMELGPRIFTFMNDSSNVAMQLQRLQMDEGRR 360
QY 361 QPRGGRDHMEALRSFLDPGSGGYSWQAHADVGHVLTGTRVTECLSLDKLEAAPEAA 420
DB 361 QPRGGRDHMEALRSFLDPGSGGYSWQAHADVGHVLTGTRVTECLSLDKLEAAPEAA 420
QY 421 LVSRALQLLAEHRFWAGVVFGLGPESSDPTHEPTDLGPGHVRIKIRMDIDVVRTNKIR 480
DB 421 LVSRALQLLAEHRFWAGVVFGLGPESSDPTHEPTDLGPGHVRIKIRMDIDVVRTNKIR 480
QY 481 DRFMDPGPAADPLTDLRYVMGGSFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYVD 540
DB 481 DRFMDPGPAADPLTDLRYVMGGSFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYVD 540
QY 541 VFLRVLSRLSPLFLTLAWIYSVTTLTKAVVREKETRLDTRAMGLSRAVLWGLWFLSCL 600
DB 541 VFLRVLSRLSPLFLTLAWIYSVTTLTKAVVREKETRLDTRAMGLSRAVLWGLWFLSCL 600
QY 601 GPFLLSAALLVLKLGILPYSHPGVVFLLAFAPAVATVTSFLLSAFFSSANLAAACG 660
DB 601 GPFLLSAALLVLKLGILPYSHPGVVFLLAFAPAVATVTSFLLSAFFSSANLAAACG 660
QY 661 GLAYFSLYLPYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEQGEGAQHNVG 720
DB 661 GLAYFSLYLPYVLCVAMRDRLPAGGRVAASLLSPVAFGFCESLALLEEQGEGAQHNVG 720
QY 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPWNFFRRSYWCGPRPP 780
DB 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPWNFFRRSYWCGPRPP 780
QY 781 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPALRGLSLDFYQGHITAPL 840
DB 781 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPALRGLSLDFYQGHITAPL 840
QY 841 GHNGAGKTTTLLSLSGLFPSPSGSAPIIGHDVRSMAAIRPHLGVCPOYNVLPDMLTVDE 900
DB 841 GHNGAGKTTTLLSLSGLFPSPSGSAPIIGHDVRSMAAIRPHLGVCPOYNVLPDMLTVDE 900
QY 901 HWFYGRKLKGLSAAVWGPEQDRLLQDVLVSKQSVQTRHLSGMGORKLSVAIAFVGGSQV 960
DB 901 HWFYGRKLKGLSAAVWGPEQDRLLQDVLVSKQSVQTRHLSGMGORKLSVAIAFVGGSQV 960
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLLDEALLGDRVAVVAGGLCCCG 1020
DB 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLLDEALLGDRVAVVAGGLCCCG 1020
QY 1021 SPLPLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKNGSQGSRVGTPLLIA 1080
DB 1021 SPLPLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSDVTRQEKNGSQGSRVGTPLLIA 1080

Db 1021 SPFLRHLGSGYVLTIVKARLPLTTNEKADTMEGSDVTRQEKKGSGQSRVGTQPQLLA 1080
Qy 1081 LVQHWVPGARLVEELPHELVLVLVPTGAGDGSFATLFPRELDTRLAELRLTGYGIDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLVPTGAGDGSFATLFPRELDTRLAELRLTGYGIDTSLE 1140
Qy 1141 EIFLKVVEECAADTDMEDSGCGHLCGTGAGLDVTLRLKMPQPTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTDMEDSGCGHLCGTGAGLDVTLRLKMPQPTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRVOGNALTRQOQLALLKRFLLARSRRGLFAQIVLPALFVGLALVFSLLIV 1260
Db 1201 QGSGPDVAVRVOGNALTRQOQLALLKRFLLARSRRGLFAQIVLPALFVGLALVFSLLIV 1260
Qy 1261 PPFCHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGLEPPVQHSSHRSFA 1320
Db 1261 PPFCHYPALRLSPMTYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGLEPPVQHSSHRSFA 1320
Qy 1321 PEVPAEVAKVLASGNWTPESPSPACQSCPGARLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Db 1321 PEVPAEVAKVLASGNWTPESPSPACQSCPGARLLPDCPAAAGGPPPPQAVTSGGEVQ 1380
Qy 1381 NLGTRNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPLSPGOELGRSVEELW 1440
Db 1381 NLGTRNLSDFLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPLSPGOELGRSVEELW 1440
Qy 1441 ALLSPLFGGALDRVLKNTAWAHSDDQSLIKIWFNNKWHSMVAFVNRASNAILRAHLP 1500
Db 1441 ALLSPLFGGALDRVLKNTAWAHSDDQSLIKIWFNNKWHSMVAFVNRASNAILRAHLP 1500
Qy 1501 PGPARHAHSITLNLHPLNLTKEQLSEALMASSVDVLVSVICVWFAMSFVPASFTVLVIEE 1560
Db 1501 PGPARHAHSITLNLHPLNLTKEQLFEALMASSVDVLVSVICVWFAMSFVPASFTVLVIEE 1560
Qy 1561 RVTRAKHLQLMGGLSPTLVWLNGLFMDNCNVLVPAQIVLFIQFQRAVYVAPANLPA 1620
Db 1561 RVTRAKHLQLMGGLSPTLVWLNGLFMDNCNVLVPAQIVLFIQFQRAVYVAPANLPA 1620
Qy 1621 LLLLLYGWSTITPLMPYASFVSPSTAYVVLTCINLFIGINGSMATFVLELFSQKLOEV 1680
Db 1621 LLLLLYGWSTITPLMPYASFVSPSTAYVVLTCINLFIGINGSMATFVLELFSQKLOEV 1680
Qy 1681 SRIILKQVFLIPFHCCLGRGLIDMVRNQAMADAFERLGRDQFOSPLRVEVVGKILLAMV 1740
Db 1681 SRIILKQVFLIPFHCCLGRGLIDMVRNQAMADAFERLGRDQFOSPLRVEVVGKILLAMV 1740
Qy 1741 GFLFLLFTLLLOHRSQLLPQPRVRSPLPLGGEDEDAVARERVVQAGTQGDVLRNLTK 1800
Db 1741 GFLFLLFTLLLOHRSQLLPQPRVRSPLPLGGEDEDAVARERVVQAGTQGDVLRNLTK 1800
Qy 1801 VYRGORMPAVDRLCLGIPPGSCFGLLVGNGAGKTSFTFVMTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRLCLGIPPGSCFGLLVGNGAGKTSFTFVMTGDTLASRGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSNYAD 1920
Db 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSNYAD 1920
Qy 1921 RPAGTYSGGNKRLKATALLVGDPAVVFLEPPTTGMPPSARRFWNSLLAVVREGSRVML 1980
Db 1921 RPAGTYSGGNKRLKATALLVGDPAVVFLEPPTTGMPPSARRFWNSLLAVVREGSRVML 1980
Qy 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAA 2040
Db 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAA 2040
Qy 2041 EFPGSSELRHAGGRRLRFLPFGRCALARVFGELAVHGAEGVDFSVQTMLEEVFLYF 2100
Db 2041 EFPGAELREARAGGRRLRFLPFGRCALARVFGELAVHGAEGVDFSVQTMLEEVFLYF 2100
Qy 2101 SKDQKQEDTEEOKEAGVGDPAFGLOHPKRVSVOLFDDPSTAEIVL 2146
Db 2101 SKDQKQEDTEEOKEAGVGDPAFGLOHPKRVSVOLFDDPSTAEIVL 2146

RESULT 3
Q9NR73
ID Q9NR73 PRELIMINARY; PRT; 2146 AA.
AC Q9NR73;
DT 01-OCT-2000 (T-REMBLrel. 15, Created)
DT 01-OCT-2000 (T-REMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-REMBLrel. 26, Last annotation update)
DE Macrophage ABC transporter.
GN Name=ABCA7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20334305; PubMed=10873640;
RA Kaminski W.E., Orso E., Diederich W., Klucken J., Drobnik W.,
RA Schmitz G.;
RT "Identification of a novel human sterol-sensitive ATP-binding cassette
RT transporter (ABCA7).";
RL Biochem. Biophys. Res. Commun. 273:532-538 (2000).
CC -|- SMITARIITY: Belongs to the ABC transporter family.
DR EMBL; AF250238; AA85794.1; -;
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0005524; F: ATP binding; TAS.
DR GO; GO:0004009; F: ATP-binding cassette (ABC) transporter acti. .; TAS.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF000005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234468 MW; 679B16EB2D75FF0D CRC64;

Query Match 99.9%; Score 11128; DB 2; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAFTQMLLLWKNMYRRQPVQLLVLLWPLFFILVAVRHSHPLEHHECHFPNKP 60
Db 1 MAFTQMLLLWKNMYRRQPVQLLVLLWPLFFILVAVRHSHPLEHHECHFPNKP 60

Qy 61 LPSAGTVPMLOGLICNVNNTCFPOLTPGEEPGRLSNFNDLSVRLADARTVLGSAHR 120
Db 61 LPSAGTVPMLOGLICNVNNTCFPOLTPGEEPGRLSNFNDLSVRLADARTVLGSAHR 120

Qy 121 TLAGLGKLIATLRAARSTAOPTKQSPLEPPMLDVALLTSLRTESLGLALQAOEPL 180
Db 121 TLAGLGKLIATLRAARSTAOPTKQSPLEPPMLDVALLTSLRTESLGLALQAOEPL 180

Qy 181 HSLLAEADLAOELLARLSVELRALLQRPRTSGPGLLELSEALCSVRGSSSTVGPSSLNW 240
Db 181 HSLLAEADLAOELLARLSVELRALLQRPRTSGPGLLELSEALCSVRGSSSTVGPSSLNW 240

Qy 241 YEASDLMELVQGPESALPDSSLSPPACSELIGALDSHPLSRLLWRLKPLILGLKLLFAPD 300
Db 241 YEASDLMELVQGPESALPDSSLSPPACSELIGALDSHPLSRLLWRLKPLILGLKLLFAPD 300

Qy 301 TPFTKRLMAQVNTTFEELTLRLDRVREVMELGPRITFTFMDSSNVAMLQRLQMDSGRR 360
Db 301 TPFTKRLMAQVNTTFEELTLRLDRVREVMELGPRITFTFMDSSNVAMLQRLQMDSGRR 360

Qy 361 QPRPGGRDHMEALRSFLDPGSGGYWQDAHADVCHLGLTGRVTECLSLDKLEAPSEAA 420
Db 361 QPRPGGRDHMEALRSFLDPGSGGYWQDAHADVCHLGLTGRVTECLSLDKLEAPSEAA 420

Qy 421 LVSRALQLLAHRAHFWAGVFLGPDSDSDPTHEPTDGLPGHVRIRKIRMDIDVVTTRNKIR 480
Db 421 LVSRALQLLAHRAHFWAGVFLGPDSDSDPTHEPTDGLPGHVRIRKIRMDIDVVTTRNKIR 480

QY 481 DRFDGPAADPLTDLRYYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQOMPPCYVDD 540
DB 481 DRFDGPAADPLTDLRYYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQOMPPCYVDD 540
QY 541 VFLRVLSRLPLPLTLAWIYSVTLVKAVVREKETRLDTRAMAGLSRAVLMLGWFLSCL 600
DB 541 VFLRVLSRLPLPLTLAWIYSVTLVKAVVREKETRLDTRAMAGLSRAVLMLGWFLSCL 600
QY 601 GFPLLSAALLVLVLGDLILPYSHPGVWFLFLAAFAVAVTQSFLLSAFFSRANLAAACG 660
DB 601 GFPLLSAALLVLVLGDLILPYSHPGVWFLFLAAFAVAVTQSFLLSAFFSRANLAAACG 660
QY 661 GLAYFSLYLPYVLCVAMRDLRDPAGGRVAASLLSPVAFGCSLALLBEGGCAQWHNVG 720
DB 661 GLAYFSLYLPYVLCVAMRDLRDPAGGRVAASLLSPVAFGCSLALLBEGGCAQWHNVG 720
QY 721 TRPTADVLSLAQVSGLLLDLDAALYGLATWYLEAVCPQYGIPEPMNPPFRSRWYCGPRPP 780
DB 721 TRPTADVLSLAQVSGLLLDLDAALYGLATWYLEAVCPQYGIPEPMNPPFRSRWYCGPRPP 780
QY 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGLSLDFYQGHITAF 840
DB 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGLSLDFYQGHITAF 840
QY 841 GHNGAGKTTTSLGSLFPFGSGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTVD 900
DB 841 GHNGAGKTTTSLGSLFPFGSGSAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTVD 900
QY 901 HWFYGRKLGSAAVVGPQDRLQDVGLVSKOSVQTRHLSGGMQRKLSVAIAFYGGSOV 960
DB 901 HWFYGRKLGSAAVVGPQDRLQDVGLVSKOSVQTRHLSGGMQRKLSVAIAFYGGSOV 960
QY 961 VILDEPTAGVDPASRRGIWELLKRYREGRTLLSTHLLDEAELLGDRVAVAGRLCCCG 1020
DB 961 VILDEPTAGVDPASRRGIWELLKRYREGRTLLSTHLLDEAELLGDRVAVAGRLCCCG 1020
QY 1021 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKKNGSGSRVGTPOLLA 1080
DB 1021 SPLFLRRHLGSGYLLTVKARLPLTTNEKADTDMEGSVDTROEKKNGSGSRVGTPOLLA 1080
QY 1081 LVQHVVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLRLAELRLTGYISDTSLE 1140
DB 1081 LVQHVVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLRLAELRLTGYISDTSLE 1140
QY 1141 EFLKVVBECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
DB 1141 EFLKVVBECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
QY 1201 QSGGPDVGRVQGWALTRQQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFSLTV 1260
DB 1201 QSGGPDVGRVQGWALTRQQLQALLKRPFLARRRRGLFAQIVLPALFVGLALVFSLTV 1260
QY 1261 PPFHYPALRLSPTMYGAQVSPFSDAPGPGARLLEALQEAAGLEPPVQHSRRFSA 1320
DB 1261 PPFHYPALRLSPTMYGAQVSPFSDAPGPGARLLEALQEAAGLEPPVQHSRRFSA 1320
QY 1321 PEVPAEVAKLASGNWTPESPSPACQCPGARRLLPDCPAAAGPPPPQAVTSGGEVQ 1380
DB 1321 PEVPAEVAKLASGNWTPESPSPACQCPGARRLLPDCPAAAGPPPPQAVTSGGEVQ 1380
QY 1381 NLTGRLNSDLFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGGRDPLGPSQELGRSVEELW 1440
DB 1381 NLTGRLNSDLFLVKTYPRLVROGLKTKKWNVEVRYGFSGLGGRDPLGPSQELGRSVEELW 1440
QY 1441 ALLSPLPGCALDRVLKNTAWAHSDDAQSLKIWNFNKGMWSWAFVNRASNAILRAHLP 1500
DB 1441 ALLSPLPGCALDRVLKNTAWAHSDDAQSLKIWNFNKGMWSWAFVNRASNAILRAHLP 1500
QY 1501 PGPARRHAHSITTLNHPNLTKQLSEALMASSVDVLVSIQVVPFMSFVPASFTLVLE 1560
DB 1501 PGPARRHAHSITTLNHPNLTKQLSEALMASSVDVLVSIQVVPFMSFVPASFTLVLE 1560

QY 1561 RVTRAKHLQMGSLPTLYLGNFLWDMCNLYLPACIVLVILFLAFOORAVVAPANLPALL 1620
DB 1561 RVTRAKHLQMGSLPTLYLGNFLWDMCNLYLPACIVLVILFLAFOORAVVAPANLPALL 1620
QY 1621 LLLLYGWSITPLMYPASFFSVPTAYVVLTCINLPIGINGSMATFVLELSDQKLQEV 1680
DB 1621 LLLLYGWSITPLMYPASFFSVPTAYVVLTCINLPIGINGSMATFVLELSDQKLQEV 1680
QY 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGRQFQSPLEWVGKLLAMVIQ 1740
DB 1681 SRLIKQVFLIFPHFCLGRGLIDMVRNOAMADAFERLGRQFQSPLEWVGKLLAMVIQ 1740
QY 1741 GFLFLFTLLQHRSQLLPQPRVRSPLLGEDEDEVARERERVQCATOGDVLVLRNLTK 1800
DB 1741 GFLFLFTLLQHRSQLLPQPRVRSPLLGEDEDEVARERERVQCATOGDVLVLRNLTK 1800
QY 1801 VTRGQMPAVDRCLGIPGECFGLLVGNAGAKTSTFRMVTGDTLASRGEAVLAGHSVAR 1860
DB 1801 VTRGQMPAVDRCLGIPGECFGLLVGNAGAKTSTFRMVTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLRCVPEAOVAGTAGSLARGLGSWYAD 1920
DB 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLRCVPEAOVAGTAGSLARGLGSWYAD 1920
QY 1921 RPAGTYSNGNKKLATALALVGDPAVVFLEDEPTTGMDPSARRFLMNSLLAVVREGRSVML 1980
DB 1921 RPAGTYSNGNKKLATALALVGDPAVVFLEDEPTTGMDPSARRFLMNSLLAVVREGRSVML 1980
QY 1981 TSHSMECEALCSRLAIWVNGFRCLGPSQHLKGRFAAGHTTLTLRVPAARSQPAAAFVAA 2040
DB 1981 TSHSMECEALCSRLAIWVNGFRCLGPSQHLKGRFAAGHTTLTLRVPAARSQPAAAFVAA 2040
QY 2041 EPPGSELREAHGRLRFQPPGRCALARVFGELAVHGAHEGVDESVSQTMLLEVLYF 2100
DB 2041 EPPGSELREAHGRLRFQPPGRCALARVFGELAVHGAHEGVDESVSQTMLLEVLYF 2100
QY 2101 SKDQCKDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAEVTL 2146
DB 2101 SKDQCKDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 4

Q96S58 PRELIMINARY; PRT; 2008 AA.
AC Q96S58;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ABCA-SSN.
GN Name=ABCA7/ABCA-SSN;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21255283; PubMed=11355874;
RA Tanaka A., Ikeda Y., Abe-Dohmae S., Arakawa R., Sadanami K.,
RA Kidera A., Nakagawa S., Nagase T., Aoki R., Kioka N., Amachi T.,
RA Yokoyama S., Ueda K.;
RT "Human ABCA1 Contains a Large Amino-Terminal Extracellular Domain
Homologous to an Epitope of Sjogren's Syndrome";
RL Biochem. Biophys. Res. Commun. 283:1019-1025(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB055390; BAB62294.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.

RESULT 5	Q7TNJ2	PRELIMINARY;	PRT; 2170 AA.	
ID	Q7TNJ2	AC	Q7TNJ2;	
DT	01-OCT-2003	(T:EMBLrel. 25, Created)		
DT	01-OCT-2003	(T:EMBLrel. 25, Last sequence update)		
DT	01-MAR-2004	(T:EMBLrel. 26, Last annotation update)		
DE	ATP-binding cassette transporter sub-family A member 7.			
GN	Name=ABCA7;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Wister; TISSUE=Platelet;			
RA	Sasaki M., Nada S., Yamaguchi A.;			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: Belongs to the ABC transporter family.			
DR	EMBL; AB097814; BAC81426.1; -.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.			
DR	GO; GO:0000166; P:nucleotide binding; IEA.			
DR	GO; GO:0006810; P:transport; IEA.			
DR	InterPro; IPR003593; AAA_ATPase.			
DR	InterPro; IPR003439; ABC transporter.			
DR	Pfam; PF00005; ABC_tran; 2.			
DR	ProDom; PD000006; ABC_transporter; 2.			
DR	SMART; SM00382; AAA; 2.			
DR	PROSITE; PS00211; ABC_TRANSPORTER_1; 1.			
DR	PROSITE; PS00893; ABC_TRANSPORTER_2; 2.			
KW	ATP-binding.			
Q	SEQUENCE 2170 AA; 237718 MW; 003C8DF70B8744CE CRC64;			
Query Match	77.6%;	Score 8652;	DB 2; Length 2170;	
Best Local Similarity	76.8%;	Pred. No. 0;		
Matches 1687;	Conservative 157;	Mismatches 274;	Indels 80; Gaps 12;	
QY	1	MAFWTQMLLLKWNFMYRRRQPVQLLVLLVPLFLFFLTVAVRSHHPLEHHECHFPNKP	60	
DB	1	MAFWTQMLLLKWNFYRRRQPIQLVLLVPLFLFFLTVAVRSHHPLEHHECHFPNKP	60	
QY	61	LPSAGTVPMQLGI CNVNNTCPQPTGPEGRSLNFNDLSVRLADARTVLGGASAH	120	
DB	61	LPSAGTVPMQLGVCNVNNSCFQFTPGKPGVLSNFKDSILIRLLADAHVLGGHSTQD	120	
QY	121	TLAGLGKLIATLRARSTHQP-----OPTKQSPLEPPMDVAELTSLIRTESLGALGOA	176	
DB	121	MLAALGKLIPLVRAGSGAWPOQESQPAKQG-----SVTELLEKILQRASLETVLGQA	173	
QY	177	QEPHLSLEAEADLAQELLARLSVELRALLORPRTGSGPLELSEALCSVRGSSVTGP	236	
DB	174	QSMRKPSDATRTVAQELLTLPSLVELRALLRRPRGSGAGLELSEALCSKSGSPGL	233	
QY	237	SLNWEASDLMEVLVGOEPESALPDSLSLSPACSELTGALDSHPLSRLLWRRLKPLILGKLL	296	
DB	234	SLNWEANQINFMGPELAPTLPDSSLSPACSEFVGALDHPVSRLLWRRLKPLILGKIL	293	
QY	297	PAPDPTFTKLAQVNRFTFEILTLLRDVREVMELGPRIFTFMNDSNVAMLQRLQMD	356	
DB	294	FAPDTNFTKLAQVNRFTFEILTLLRDVREVMELGPRIFTFMNDSNVAMLQRLQMD	353	
QY	357	EGRPQPRGGRDHMEALRSLDPGSGGYSMQDADHVGHLVGTIGRVTCLSLDKLEAP	416	
DB	354	TGQQQTTPGQQLSAIRDFLDPGRGRYVNOEADHMGRLAEIIGQILCVCSLDKLEAVP	413	
QY	417	SEAAVSRALQLLAHRFWAGVVFVFLGPDSSDPTHEHTP-DLGPCHVRKIRMDIDVVTR	475	
DB	414	SEAAVSRALQLLAHRFWAGVVFVFLGPDSSDPTHEHTP-DLGPCHVRKIRMDIDVVTR	473	
QY	476	TNKIRDFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYP	535	

Qy 1569 QLMGSLPTLYWLGWFLWDMCNVLPACIVVLIIFLAFOORAYVAPANLPALLILLILLYGW 1628
 Db 1594 QLVSGLPQTLWLGWFLWDMCNVLPACIVVLIIFLAFOORAYVAPENLPALLILLILLYGW 1653
 Qy 1629 SITPLMPASFFFSVPSTAYVVLTCINLIFIGINSMATFVLELPSDOKLQEVSRILKQVF 1688
 Db 1654 SITPLMPASFFFSVPSTAYVVLTCINLIFIGINSMATFVLELPSDOKLQEVSRILKQVF 1713
 Qy 1689 LIFPHFCLGRGLIDMVRNQAMADAFERLGDQFOSPLRWEVVGKLLAMVIGQFLFLFT 1748
 Db 1714 LIFPHFCLGRGLIDMVRNQAMADAFERLGDQFOSPLRWDIIGKLLAMVAQGFLLIT 1773
 Qy 1749 LLLQHRQLLPQVRVSLPLGDEDEVARERERVVGATQGDVILVRLNLTKVYRGQMP 1808
 Db 1774 LLLQHRNLLPQPKSRLLPPPLGDEDEVDVREVRVTKGQGDVILVRLDLTKVYRGQSP 1833
 Qy 1809 AVDRCLGIPGECFGLLGVNGACKTSTFRMVTGDTILASRGEAVLAGHSVAREPSAAHLS 1868
 Db 1834 AVDRCLGIPGECFGLLGVNGACKTSTFRMVTGDTILPSSGEAVLAGHVAQEPSAAHRS 1893
 Qy 1869 MGYCPQSDAIFPELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLSWADRPAGTYSG 1928
 Db 1894 MGYCPQSDAIFPELLTGREHLELLARLGVPEAQVAQTAGSGLARLGLPSVADRPAGTYSG 1953
 Qy 1929 GNKRKLATALVGDPAVFLDEPTTGMDPSARFLWNLALLVREGSVWLTSHSMEEC 1988
 Db 1954 GNKRKLATALVGDPAVFLDEPTTGMDPSARFLWNLALLVREGSVVLTSHSMEEC 2013
 Qy 1989 EALCSRLAIVMNGFRCLGSPQHLKGRFAAGHTLTLRPAARSOPAAAFVAAEPGSELR 2048
 Db 2014 EALCTRILAIWNGFRCLGSGAHLKSRFGAGHTLTLRVPDPQPPALAFIVTFPPDAELR 2073
 Qy 2049 EAHGRLRFLQPPGRCRCALARVFGELAVHGAHGVEDFSVSQTMLEEVLYFSKQDQKDE 2108
 Db 2074 EVHGSRLRFLQPPGCGCTLARVPRELAQAQKAHGVEDFSVSQTMLEEVLYFSKQDQKEE 2133
 Qy 2109 DTEQKQKAGVGDPAVGLQHPKRVSRFLQDLPDSTAEVTL 2146
 Db 2134 EGSGQETETREVS-TPGLQHPKRVSRFLQDLPDSTAEVTL 2170

RESULT 6
 Q91V24 PRELIMINARY; PRT; 2159 AA.
 AC Q91V24;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ATP-binding cassette transporter sub-family A member 7.
 GN Name=Abca7;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O1129, and DBA/2;
 RX MEDLINE=21328888; PubMed=11435699;
 RA Broccardo C., Osorio J., Luciani M.-P., Schriml L.M., Prades C.,
 RA Shulenin S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,
 RA Mattel M.G., Dean M., Deneffe P., Chimini G.;
 RT "Comparative analysis of the promoter structure and genomic
 RT organization of the human and mouse Abca7 gene encoding a novel ABCA
 RT transporter";
 RL Cytochrome. Cell Genet. 92:264-270(2001).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AF287142; AAK56863.1; -;
 DR EMBL; AF287141; AAK56862.1; -;
 DR HSP; P58301; 1F2U.
 DR MGD; MGI:1351646; Abca7.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti...; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR InterPro; IPR02016; Peroxidase.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 KW ATP-binding
 SQ SEQUENCE 2159 AA; 236981 MW; C2B3FE0D8B822B CRC64;
 Query Match 77.2%; Score 8600.5; DB 2; Length 2159;
 Best Local Similarity 76.6%; Pred. No. 0;
 Matches 1676; Conservative 162; Mismatches 280; Indels 69; Gaps 11;
 Qy 1 MATWTQLMLLLWKNFMVRRROPVQLLVLLWPLFLFFILVAVRHSHPLEHHECHFPNKP 60
 Db 1 MALGTQLMLLLWKNYTYRRRQPIQLLVLLWPLFLFFILVAVRHSHPLEHHECHFPNKP 60
 Qy 61 LPSAGTVPWLQGLICNVNNTCFPOLTPGEPGRSLSNFNDLSVRLADARTVLGGASAH 120
 Db 61 LPSAGTVPWLQGLICNVNNSCFQHPPTGEPKPGVLSNFKDSLIRLLADRTFVLGGHIQD 120
 Qy 121 TLAGLGKLIATLRAASTAQP-----OPTKOSPLEPPMLDVAELLTSLRRTESLALGQA 176
 Db 121 MLDALGKLIPLVRAVGAGARPQSDQPTSG-----SVTKLEKILQRLADPVLGQA 173
 Qy 177 QEPHLSLEAEADLAELLALRSILVELRALLQRPRTSGSPLELLSEALCSVRGSSVTGP 236
 Db 174 QDSMRKFSDAIRLAQELLTLPSLMELRALLRRPRGSAGLELVSEALCSKGPSSPGL 233
 Qy 237 SLNMYEASDLIMELVGPESALPDSLSLSPACSELIGALDSHPISRLLRLLKPLILGKLL 296
 Db 234 SLNMYEANQLNFMGPEVAPALPDNSLSPACSEFVGTLDHPVSRLLRRLKPLILGKIL 293
 Qy 297 FAPDTPFTRKLMQAVNRTFEEELTLRDVREWEMLGPRIFTFMNDSSNVAMLRQLQM 356
 Db 294 FAPDTNTRKLMQAVNQTPFEELALLRDLHELWGLVGPQIFNFMNDSTNVAMLRQLLDVGG 353
 Qy 357 EGRROPQGRDHWELRSLFDPSGGYSGWQDAHADVHGLVGLTRVTECLSDKLEAAP 416
 Db 354 TGORQQTTPRAQKLEAKDFDPSRGYSWREAHDMGRLAGILQGMCEVSLDKLEAVP 413
 Qy 417 SEALVSRALQLLAERFWAGVYFLGPDSSDTEHPTDLPGLGHVRIKIRMDIDVVTRT 476
 Db 414 SEALVSRALLELGERRLWAGIVFLSPEHPLDPSSELSPPALSPGHLPFKIRMDIDVVTRT 473
 Qy 477 NKIRDRFMDPGPAADPLTDLRYVYMGGFVYLQDLIVERAAVRVLSGANPRAGLYLQMPYPC 536
 Db 474 NKIRDFMDPGSADPFMDLRYVYMGGFVYLQDLLEQAARVYVGGNSRTGLYLQMPHPC 533
 Qy 537 YVDDVFLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDTRMANGLSRAVLWLGWF 596
 Db 534 YVDDVFLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDTRMANGLSRAVLWLGWF 593
 Qy 597 LSCLGPFLLSAALLVLVGLDILPYSHPGVWFLFAAPAVATVTSQFLLSAPFSRANLA 656
 Db 594 LSCLGPFLLSAALLVLVGLGNILPYSHPVVIFLFAAFVATVATQSFLLSAFESRANLA 653
 Qy 657 AACGGLAYFSLYLPYVLCVAVNRDLRPPAGGRVAASLLSPVAFGCGCSLALLEBQGAQW 716
 Db 654 AACGGLAYFALYLPYVLCVAVNRERHLGGLAASLLSPVAFGCGCSLALLEBQGAQW 713
 Qy 717 HNVGTRPTADVSLAQVSGLLILLDAALYGLATWYLEAVCPQYGIPEPWNFPFRSVMCG 776
 Db 714 HNLGTGAEDVFSLAQVSAFLLLDVAYIGLAWYLEAVCPQYGIPEPWNFPFRSVMCG 773
 Qy 777 PRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQAPLRGLSLDIFYGHI 836

Db 774 PGPSSVLAAPADQPKVLEPPGLVPGVSIIRGLKXHFRCPOPALQGLNLDPEYGH 833
Qy 837 TAFLGHGAGKTTTILSILSGLFPFPPSGGSAF ILGHDRVRSMAAIRPHLGVCQYNVLFDM 896
Db 834 TAFLGHGAGKTTTILSILSGLFPFPPSGGSAF ILGHDRVRSMAAIRPHLGVCQYNVLFDM 893
Qy 897 TVDEHVWYGRKLGSAVWVGEQDRLLQDVLGSKSVQTRHLSGGHQRKLSVAIAFVG 956
Db 894 TVEEHVWYGRKLGSAVWVGEQDRLLQDVLGSKSVQTRHLSGGHQRKLSVAIAFVG 953
Qy 957 GSOVILDEPTAGVDPASRRGIWELLLXYREGRTILSTHLDDEALLGDRVAVAGRL 1016
Db 954 GSRVIMDEPTAGVDPASRRGIWELLLXYREGRTILSTHLDDEALLGDRVAVAGRL 1013
Qy 1017 CCCGSPFLRRHLSGYVLTIVKARLPLTITNE-KADTMEGSDVPRQKNGSQSGS---- 1071
Db 1014 CCCGSPFLRRHLSGYVLTIVKARLPLTITNE-KADTMEGSDVPRQKNGSQSGS---- 1068
Qy 1072 -----RVGTPQALLVQHVWPGARLVEELPHELVLVLPVYTGADH 1110
Db 1069 APTRGTSKSNQAPAPGAVPTTPSTARILELVQHVPGQAVLELPHELLVLVYPAGALD 1128
Qy 1111 GSPATLRELDRTRLAELRLTGVISDTSLEELFLKVEECRADTDMEDGSCGHLCTGIA 1170
Db 1129 GSFAMVFOELDQQLLEGLTGYISDITNLEELFLKVEED--AHREGGDSRPLHLRT--- 1183
Qy 1171 GLDVTLRKMPPOETALENGEPAGSAPETDQSGPDVGRVQGWALTQOQLQALLKRFL 1230
Db 1184 ---CTPQPTGPEASVLENGELA-----PQGLAANA-AQVQGWLTTCQRLALLKRFL 1233
Qy 1231 LARRSRRLGFAQIVLPALFVGLALVFLTPFPFGHYPALRLSPTMYGAQVFFSBDAPGD 1290
Db 1234 LARRSRRLGFAQIVLPALFVGLALVFLTPFPFGHYPALRLSPTMYGAQVFFSBDAPGD 1293
Qy 1291 PGRARLLALQEALEPPVVOHSHR-----FSAPEVPAEVAKVLAGSWTPE 1339
Db 1294 PNRMKLLLEALGEAGLQPSMDKARGSECTHSLACYFTTVEVPDPAVASILASGNWTP 1353
Qy 1340 SPSPACQCSQPCARLLPDCPAAAGPPPPQAVTSGGEVQNLTRGNLSDFLVKTYPRLV 1399
Db 1354 SPSPACQCSQPCARLLPDCPAGAGPPPPQAVAGGEVQNLTRGNLSDFLVKTYPSLV 1413
Qy 1400 RQGLTKKKNVRYGGFSLGGRDPGLPSGQELGRSVEELWALLSPLFCGALDRVLKNLT 1459
Db 1414 RGLTKKKNVRYGGFSLGGRDPGLTGEVVRTLAERALLSPQGNALDRILNLT 1473
Qy 1460 AWASLDAQDSLIKIFNNKWHMSVAFVNRASNAILRAHLPPGPARHAHSITTLNHLNL 1519
Db 1474 QWALGLDARNSLIKIFNNKWHAMVAFVNRANGLLHALLSPGVRHAHSITTLNHLNL 1533
Qy 1520 TKEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVLEERTVTEAKHLQMLGGLSPTLY 1579
Db 1534 TKEQLSEALMASSVDVLVSVICVVFAMSFVPASFTLVLEERTVTEAKHLQMLGGLSPTLY 1593
Qy 1580 WLGNFLDMCNVLPACIVLIFLAFQORAVYAPANLPALELLLLLYGWSITPLMPYASF 1639
Db 1594 WLGNFLDMCNVLPACIVLIFLAFQORAVYAPANLPALELLLLLYGWSITPLMPYASF 1653
Qy 1640 FFSVSTAYVVLTCINLFIGINGSWATVLELFSQKLQEVSRILKQVLFIPPHFCLG 1699
Db 1654 FFSVSTAYVVLTCINLFIGINGSWATVLELFSQKLQEVSRILKQVLFIPPHFCLG 1713
Qy 1700 LIDMYRNQAMADAFERLQDRQFQSLRVEVVGKLLAMVIOGPLELLTLLLOHRSQLLP 1759
Db 1714 LIDMYRNQAMADAFERLQDRQFQSLRVEVVGKLLAMVIOGPLELLTLLLOHRSQLLP 1773
Qy 1760 QPRVRSLLPGLBEDEVDARERERVQATQGVDLVRLNLTKVYRGORMPAVDRLCLGIPP 1819
Db 1774 QSKPRLPPLGEBEDEDVAQERERVTKGATQGVDLVRLNLTKVYRGORMPAVDRLCLGIPP 1833
Qy 1820 GECFGLGVNGAGKTSTFRMTVGTDLASRGEAVLAGHSVAREPSAHLGSMGYCQSDAIF 1879
Db 1834 GECFGLGVNGAGKTSTFRMTVGTDLASRGEAVLAGHSVAREPSAHLGSMGYCQSDAIF 1893

Qy 1880 ELLTGREHLELLARLGRVPEAQVTAQTAGSLARLGLSWYADRPACTYSGGNKRKLATALA 1939
Db 1894 DLLTGREHLELLARLGRVPEAQVTAQTAGSLARLGLSWYADRPACTYSGGNKRKLATALA 1953
Qy 1940 LVGDPVAVFLDEPTTGMPSARRFLWNSLLAVVREGRSVMTLTSMSERCEALCSRLAIMV 1999
Db 1954 LVGDPVAVFLDEPTTGMPSARRFLWNSLLAVVREGRSVMTLTSMSERCEALCSRLAIMV 2013
Qy 2000 NGRFCLGSPHOLKGRFAAGHTTLRVPAARSQPAAPAAAFVAAEFPGSELREAHGGLRFL 2059
Db 2014 NGRFCLGSSHLKGRFGAGHTTLRVPPDQPEFAIAFIRITFPGAELREHVHGLRFL 2073
Qy 2060 PPGRCALARYFGLAVHGAEGHVEDFSVSQTMLEEVFLYFSKQGDDETEEQKAGVG 2119
Db 2074 PPGRCGLTRVFERELAAQGRAHGVDFSVSQTMLEEVFLYFSKQGBEESRSRQ-EABEE 2132
Qy 2120 VDPAGLOHPKRVSQFLDDPSTAEVTL 2146
Db 2133 EVSKPGRQHPKRVSRFLEDPSSVETMI 2159
RESULT 7
QyUVV4 PRELIMINARY; PRT; 2260 AA.
AC Q8UVV4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ATP-binding cassette transporter 1.
GN Name=ABCA1;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Attie A.D., Brooks-Wilson A.R., Walker D., McManus B.,
Gray-Keller M.P., MacDonald M.L.E., Roomp K., Tebon A., Zhang L.-H.,
Mulligan J., Seneen C., Bitgood J.J., Cook M.E., Kastelein J.J.P.,
Hayden M.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: Belongs to the ABC transporter family.
EMBL; AF362377; AAL56247.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA-ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2260 AA; 254070 MW; 19D137F343F98662 CRC64;
Query Match 52.2%; Score 5817; DB 2; Length 2260;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1151; Conservative 369; Mismatches 594; Indels 174; Gaps 25;
Qy 1 MAFWTQLMLLWKNFMYRRRRQVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MAFWTQLMLLWKNFMYRRRRQVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKA 60
Qy 61 LPSAGTVPWLOGLCNVNTCFPQITPCGEEGRSLNENDSLVSLLDARTVLGASAH 120
Db 61 MPSAGTLPWIOGLIICNANNPCFRYPTPGESPGVGNFNASIVSRFLSDAKRLLYSQDT 120
Qy 121 TLGLGLKIATLRAA----- 135

Db	121	SIKDQVQVLAKRLKGLNGSSGLDLKLRDLFDVNETFSDLRHNVSMPSSAVEELLDAEVLN	180
Qy	136	-----RSTAQPQFTKOSPLEPPMLDVAEL-----	159
Db	181	QKVIVGYRIQLRDLNCSALSEFLTQNRSVAMDSAEFLCTLPKETLHAAELAFRANLN	240
Qy	160	-LTSLLRTESLIGALGOAQOAPLHSLLEAAEDLAQELLALRSVLREL---ALLQRPRTSG	215
Db	241	PLRPLQREIIFNSSLRDLSETVEALRDSLGKVKELLSMKSWDMRQEVMLFTVNASNS	300
Qy	216	PLEL---LSBALCSVRGSPSTVGPSLWNWYBASDLMELVG---QEPSALPDSSLSPACS	268
Db	301	STQIYQAVSRIVCGHPEGGGLKIKSLWNVEDNNYKALFCGNSTEDDDVTNFVDNSTTFCYN	360
Qy	269	ELIGALDQHPRLLLWRLLPLIIGKLLFADDPFTKLAQVNRTEELTLRLDVEVW	328
Db	361	ELMKNLESSPLSRIWALPKLLIIGKVLYPDTPAIRKINAEVNRTPQELIGVRDLGGMW	420
Qy	329	EMLGPRITFTFNDSSNVAMLQRLQ-----MODEGR---RPRPGGRDHM	370
Db	421	EETSPKLTWTWESSQOEMDLITLLKSKALWDLHLPASNWTVEDVARFLSKHP-----E	475
Qy	371	EALRSFLDPGSGGYSDAHADVGHVLVGTGLGRVTECLSDKLEAAPSEAAALVSRALQLLA	430
Db	476	EADNGMV-----YTWVDAFNETDRAIQTISRFMECVNLDKLEPVA TEVRLINKSLELLD	529
Qy	431	EHRFWAGVFLPEBDDSDPTEHPTDPLGPHGVRIKIRMDIDVTRTNKIKDRFWDPGPAA	490
Db	530	ERRFWAGVFP--TEIAPNSTELPQ-----HVYKIRKIMDINVERTNKIKDGYWDPGRA	581
Qy	491	DPLTDLYVMGGFVYLQDLVERAAVRVLSGANPRAGLYLOOMPYPVCVDDVFLRVLRSLS	550
Db	582	DPEDMRVVMGGFTYLDVVEQALIRVQTGEKKTGYVVOOMPYPVCVDDIPLRMBGRSM	641
Qy	551	PLFUTLAWIYSVTLTIVKAVVREKETRLRDTVRAMGLSRAVLWLGWFLUSCLGPFLLSALL	610
Db	642	PLPMTLAWIYSVAIIKGIVVEKARLKETMRIMGLDNGILWLISWFTSSLIPLMSAGLL	701
Qy	611	VIVLKLGDILPYSHPGVVFPLFAFAVATVYQSPLLSAFSPSRANLAACCGLAYFSLYLP	670
Db	702	VLLKMGNTLLPYSDPSVVFVFLSIFGVITLIQCFLISTVFSRANLAACCGIYVFTLYLP	761
Qy	671	YVLCVAMRDLRDPAGGRVAASLLSPVAFGFGCESLALLEQEGEAOHNVGTRP--TADVFS	729
Db	762	YVLCVAMQDYVSLSKLFASLLSPVAFGFGCEYFALFEEQGVQVQWQDNFESPLEEDGFS	821
Qy	730	LAQVSGLLLLDAALYGHATWYLEAVCPQYXGIEPEPMNFPFRRSYWCGRPPKSPACOPTP	789
Db	822	ITTSAVNMVFDTPLYGVMTWYIESVPFGYIGIPRPWYFPFTKSYWFGCESDQRHLHPDQ	881
Qy	790	LDP--KVLVEAPGLSPGVSRSLKRPFGSPQPALRGLSLDFVQGHITAFLHGNGAKT	848
Db	882	KGPFSEVCKEPEMHLSLGVSIONLVKYVRDGKVAVDGLTLNFYEGQITSFGLHNGAKT	941
Qy	849	TTLSILSGLPPPSGGSFILGHDIRSSMAAIRPHLGYVCPVNVLFDMTLVDEHVWFYGR	908
Db	942	TTMSILTLGPPPTGTAFLIGKDIRSELSTIRQNLGVCPQHNVLFDLTLVEEHWFYARL	1001
Qy	909	KGLSAAVVGPEQDRLLQDVLGVSQKSVQTRHLSGGMOKLSVAIAFYGGSQVVLDEPTA	968
Db	1002	KGUPEKKVKEEMEQAMDVGLPHKLKARTSKLSGMQKLSVALAFVGGSKVVLDEPTA	1061
Qy	969	GVDPASRRGTWELLKVKREGBRTLILSTHHLDDEALLGDRVANVAGGRLCGCCGSPFLURRH	1028
Db	1062	GVDPSRRGIWELLKVKROGRTIILSTHHMDEADILGRIAIIISHGKLCVCGSSLFKNQ	1121
Qy	1029	LGSGYILTAKARLPLTTN-----EKADTDMEGSVDTROEKKNGSQGSRVCTP	1076
Db	1122	LGTYILTUVKQDVSSLSRCNSSSTVSYLKDDSVSQSSSDGLGSDHESDTLITIDVS	1181
Qy	1077	QLLALVOHVWPGARLVEELPHELVLVLPYTCANDGSPATILFRELDTRLAELRTGYGID	1136

1182	Db	1182	AINSLITKHWPEARLVEDI	GHEITVYLPYKAAKEGAPVELFHEIDDRLSDLGSISSY	1241					
1137	Qy	1137	TSLEIEFLKVVECAADTMEDGSCG	HOHLCTGIAGLDVTILRLKMPPOETALE--NGEPAG	1194					
	Db		1137	TSLEIEFLKVVECAADTMEDGSCG	HOHLCTGIAGLDVTILRLKMPPOETALE--NGEPAG	1194				
1242	Qy	1242	TTLEIEFLKVADSGVDAET	SDGTLPARNRRAFG--DRQSLRPFTEDDAPDNDSIDP	1300					
	Db		1242	TTLEIEFLKVADSGVDAET	SDGTLPARNRRAFG--DRQSLRPFTEDDAPDNDSIDP	1300				
1195	Qy	1195	SAPETDQSGPDVAG--	RVQGWALTROOQALLLKRFLLARRRRGLFAQIIVLPALFVGL	1252					
	Db		1195	SAPETDQSGPDVAG--	RVQGWALTROOQALLLKRFLLARRRRGLFAQIIVLPALFVGL	1252				
1301	Qy	1301	ESRETDLLSGMDGKGYOMK	GKLSQOQFWALLWKRLLI	AKRSKGFPAQIVLPAVFVCI	1360				
	Db		1301	ESRETDLLSGMDGKGYOMK	GKLSQOQFWALLWKRLLI	AKRSKGFPAQIVLPAVFVCI	1360			
1253	Qy	1253	ALVPFSLVPPGHPYALRLSPT	WYGAQVSPFSEDA	PGDPCRARLLEALLOEAG-----	1305				
	Db		1253	ALVPFSLVPPGHPYALRLSPT	WYGAQVSPFSEDA	PGDPCRARLLEALLOEAG-----	1305			
1361	Qy	1361	ALMFSLLVPPFGKYP	SELOQWYDEQYTFISNDAP	EDAGTQKLLDALLNKPGFGTRCMQ	1420				
	Db		1361	ALMFSLLVPPFGKYP	SELOQWYDEQYTFISNDAP	EDAGTQKLLDALLNKPGFGTRCMQ	1420			
1306	Qy	1306	--LEEBPVPQHSHRRF	SAPVAEAKVTASGNWTPES	SPACQSPGARRLLLPDCPAA	1362				
	Db		1306	--LEEBPVPQHSHRRF	SAPVAEAKVTASGNWTPES	SPACQSPGARRLLLPDCPAA	1362			
1421	Qy	1421	GHSIPDTPCTVGQKEW	TASVDSVLEIL--RGWNSMEN	SPSPCECSNEKIKMMLPVCPPG	1479				
	Db		1421	GHSIPDTPCTVGQKEW	TASVDSVLEIL--RGWNSMEN	SPSPCECSNEKIKMMLPVCPPG	1479			
1363	Qy	1363	AGGPPPPQAVTSGSEV	QNLITGRNLSDFLVKTYT	PRLVROGLKTKKWVNEVRYGGFS	1422				
	Db		1363	AGGPPPPQAVTSGSEV	QNLITGRNLSDFLVKTYT	PRLVROGLKTKKWVNEVRYGGFS	1422			
1480	Qy	1480	AGGLPPQREQD	TADILQNLITGRNISD	YLVKTYAQIIGKSLKNKIWN	VEFRYGGFSLGAR	1539			
	Db		1480	AGGLPPQREQD	TADILQNLITGRNISD	YLVKTYAQIIGKSLKNKIWN	VEFRYGGFSLGAR	1539		
1423	Qy	1423	DPG--LPSGQELGRSV	BEELWALLSPLPGCALDR	VLKNI	TAWAHS	DAQDSLKIWNKNGWH	1481		
	Db		1423	DPG--LPSGQELGRSV	BEELWALLSPLPGCALDR	VLKNI	TAWAHS	DAQDSLKIWNKNGWH	1481	
1540	Qy	1540	SSHVLPPSNEVTAIKOV	KKILBLAQSSGDRFLNN	LASFPMKGLDTKNVYKWN	FNKNGWH	1599			
	Db		1540	SSHVLPPSNEVTAIKOV	KKILBLAQSSGDRFLNN	LASFPMKGLDTKNVYKWN	FNKNGWH	1599		
1482	Qy	1482	SMVAFVNRASNAITLRAH	LPGPAPRAH	SIITL	NHPLNLTKEQLSEALMASS	VDVLVSIC	1541		
	Db		1482	SMVAFVNRASNAITLRAH	LPGPAPRAH	SIITL	NHPLNLTKEQLSEALMASS	VDVLVSIC	1541	
1600	Qy	1600	AISFLVNIINAILRAN	LQGNKPSAYGII	TA	FHPLNLTQQQLSEVALMTS	VDVLVSIC	1659		
	Db		1600	AISFLVNIINAILRAN	LQGNKPSAYGII	TA	FHPLNLTQQQLSEVALMTS	VDVLVSIC	1659	
1542	Qy	1542	VVFAMSVFPASFTVL	IEEVRTKAKHLQMG	LSPTLYWLG	NFLWDMCNLYLPACI	IVVLI	1601		
	Db		1542	VVFAMSVFPASFTVL	IEEVRTKAKHLQMG	LSPTLYWLG	NFLWDMCNLYLPACI	IVVLI	1601	
1660	Qy	1660	VIFAMSFPASFVVFL	IQERSKAKHLQFISG	VKPYIWL	ANFWMDCNVIYPATL	VII	1719		
	Db		1660	VIFAMSFPASFVVFL	IQERSKAKHLQFISG	VKPYIWL	ANFWMDCNVIYPATL	VII	1719	
1502	Qy	1502	FLAQQRAYAPANL	PALLLLLLLXWSIT	PLMYPASFFFS	VSTAYVTL	CNLFLGIN	1661		
	Db		1502	FLAQQRAYAPANL	PALLLLLLLXWSIT	PLMYPASFFFS	VSTAYVTL	CNLFLGIN	1661	
1720	Qy	1720	FICFQQKSYYS	SSNLPVALLLLYG	MSITPLMYPASFV	KIPSTAYVTL	SVNLFGIN	1779		
	Db		1720	FICFQQKSYYS	SSNLPVALLLLYG	MSITPLMYPASFV	KIPSTAYVTL	SVNLFGIN	1779	
1662	Qy	1662	GSMATFVLELFDQK	QVESRILKOVFLI	PPHFC	LGRLIDMYRNQAMADAF	ERLGRQF	1721		
	Db		1662	GSMATFVLELFDQK	QVESRILKOVFLI	PPHFC	LGRLIDMYRNQAMADAF	ERLGRQF	1721	
1780	Qy	1780	GSVATFVLELFTNN	KLNNIDILKS	VFLI	PPHFC	LGRLIDMYRNQAMADAL	ERFGENRF	1839	
	Db		1780	GSVATFVLELFTNN	KLNNIDILKS	VFLI	PPHFC	LGRLIDMYRNQAMADAL	ERFGENRF	1839
1722	Qy	1722	QSPLEWVWGKLIAM	VIQGFLEFLTLL	LQHR	SQLLPOPRVRS	LPGLLEDEED	VARERE	1781	
	Db		1722	QSPLEWVWGKLIAM	VIQGFLEFLTLL	LQHR	SQLLPOPRVRS	LPGLLEDEED	VARERE	1781
1840	Qy	1840	VSPLSMDLVGRNL	FAMAVEGVVFL	ITVL	IYRFFIKPR	PVYAKLPVN	DEDEEDVNRERQ	1899	
	Db		1840	VSPLSMDLVGRNL	FAMAVEGVVFL	ITVL	IYRFFIKPR	PVYAKLPVN	DEDEEDVNRERQ	1899
1782	Qy	1782	RVVOGATQGDVL	VLRLNLT	KVYRGOR	MPADRLCLGIP	PGCE	FCGLLVNGV	GAGKYSTF	1841
	Db		1782	RVVOGATQGDVL</						

QY 161 -----TSLLRTESTGLALGOAEPHLHSLEAAEDLAQELLALRSLSVELR---ALLQPRG 212
Db 241 TKLNSTSHLPTQHILA-----EATTVLLSLGSLAQELFSTKSWSMQREVWPLTNVNS 293
QY 213 TSGPLEL---LSEALCSVRGPSSVGPSSLNWEASDLMEVLG---DEPESALPDSLSLP 265
Db 294 SSSSTQIYQAVSRVCGHPGGGLKIKSLNWEYDNNYKALFGGNNTEEDVDTFDYNSTTP 353
QY 266 ACSLIGALDHPHLSRLRLKPLILGKLFAPDPTFTKRLMAQVNRTEELTLRLDVR 325
Db 354 YCNLDMKNLESSPLSRIIWKALKPLLVKILYTPDTPATQVMAEVNKTQELAVFHDLE 413
QY 326 EVNEMLGPRIFTFMNDSNVAMLQRLLOMQ--DEGREQPRPGGRDHMEALRSEL----- 377
Db 414 GMWEELSPOIWTFMENSQEMDLVRLTLLDSRQNDQFWEQKLDGLDWTQADIMAFIAKNPED 473
QY 378 --DPSGSGYSQWDAHADVGHVGLGRVTECLSLDKLEAAPSEAAALVSRALQILAEHRFW 435
Db 474 VQSPNGSVIYWRFAFNETNQAIQIISRPMECVNLKLEPIPTVRLINKSMELLDERKFW 533
QY 436 AGVFLGPEDSSDPTHTPDLG--PGHVRKIRMDIDVYTRTNKIRDRFMDPGAADPL 493
Db 534 AGIVFTG-----ITPDSVELPHVKYKIRMDIDNVERTNKIKDGYWDPGPRADPF 583
QY 494 TDLRYVNGGFIYLODIVERAARVULSGANPRAGLYLOOMPYPYCVVDDVFLVLSRSLPLF 553
Db 584 EDMRYVNGGFAYLODVVEQAIIRVLTSEKKTGYVVOOMPYPYCVVDDIFLRVMSRSMPLF 643
QY 554 LTLAWIYSVLTWKAVVREKETRLDRDTRAMGLSRAVLWLGWFLSCIGLPGPLLSAALVLV 613
Db 644 MTLAWIYSVAVIKSIYVEKEARLKETWRINGLNGILMFSWFVSSILPLVLSAGLLVVI 703
QY 614 LKGLDILPYSHPGVVFLFLAFAVATVYQSPFLSAPFSRANLAAACGLAYFSLYLPYL 673
Db 704 LKGNLLPYSDPSVVFVFLSVFAMVTILQCFLIETLESANLAAACGGIIFYTLPLVYL 763
QY 674 CVAWDRPLPAGGRVAASILLSPVAFGFCESLALLEEGEGEAOWHNGTRPT-ADVFSLAQ 732
Db 764 CVAWQDYVGFSGIKIFASLLSPVAFGFCVFPALFEEGGIYQWQNDLPFSPVEEDGFNLTT 823
QY 733 VSGLLLLDAALYGLATWYLEAVCPQYGIPEPMNPFRRSYWCQPRPKGPAPCPTPLD- 791
Db 824 AVSMMLFDTLYGWYTWIYEAIVFGQYGIPEWYFCTKSYWFGEEIDEKSHPGSSQKV 883
QY 792 PKVLVEBAPPLSGVSVRSLEKFPSPQALRGLSLDFYQGHITAFPLHGNGAGKTTTL 851
Db 884 SEICWEEBETHRLGVSIQNLVKYRDGMKVAVDGLALNFVEGQITSFLHGNGAGKTTM 943
QY 852 SILSGLPPSGGSFALIGHDVRSMAAIRPHLGYCQYVNLFDMLTVDEHWFYGRLLKL 911
Db 944 SILTGLFPPTSGTAYIILGKDIRSEMSIRQNLGVCQPHNLFDMLTVDEHWFYGRLLKL 1003
QY 912 SAAVVGPEQDRLLQDVGL-VSKQSVQTRHLSSGGMQRKLSVAIAFVSGSVQVILDEPTAGV 970
Db 1004 SEKVAKMEQMALDVGLPSPKLSKSTQSLSGGMQRKLSVALAFVSGSVKVVILDEPTAGV 1063
QY 971 DPASRRGIWELLKYREGRTILILSTHLDLAEALGDRVAVVAGGRCLCCGSPPLFLRHLG 1030
Db 1064 DPYSRGIWELLKYRQRTIILSTHMDLADILGDRIAIISHGKLCVGSLSFLKNQLG 1123
QY 1031 SGYILTIVKARLPLTTN-----EKADTDMEGSVDTREKKNKSGSGSVRGVTPOL 1078
Db 1124 TGYILTIVKDDVESLSLSCRNSSSTVSLCKEDSVSSQSSDAGLGSDHESDTLTIDVSAI 1183
QY 1079 LALVOHWVPGARLVEELPHVLVLPYTGADHSGFATLFRLEDTRLAELRLTGYSIDTS 1138
Db 1184 SNLRKHVSEARLVEDIGHETVYLPYEAKEGAFVELFHEIDRLDGLDGLSSIGSETT 1243
QY 1139 LEEIFLKVVECAADTDMEDG-----SCGHLCTGIGAGLDVTLRLKMPQ 1183
Db 1244 LEEIFLKVAESGVDAETSDGTLPARRNRRAFGDKQSC-LHPFTEDDAVD-----PN 1294
QY 1184 ETALENGEPAGSAPETDOGSGPDVG--RVQGWALTQOQLALLKRFLLARRSRGLFA 1241

Db 1295 DSDID-----PESRETDLLSGMDKGSYQLKGWKLTCQQFVALLWKRLLIARSRGFFA 1349
QY 1242 QIVLPALFVGLALVFSLIVPPFGHYPALRLSPMYGAQVFFSEDPADGPRARLLEALL 1301
Db 1350 QIVLPAVFCIALVFSLIVPPFGKYPSELQPMWYNEQVTFVSNDAPEMDGTQELLNALT 1409
QY 1302 QEAG-----LSEPPVQSHSHRFSAPAEVAPAEVAKVLASGNWTPESPSPACQSQPG 1351
Db 1410 KDPFGFTRCMEGNPIPTDPCAGEEDWTISPVQSIQVDFQNGNWTMKNPSPACQSSDK 1469
QY 1352 ARLLLPDCPAAAGGPPPPQAQVAVTSGSEVQNLTCRNLSDFLVKTYPRLVRKLTQKKWNE 1411
Db 1470 IKKMLPVCPPGAGLPPQPKQKTADILQNLTCRNLSDILVKTYPVQIIAKSLKNKWNE 1529
QY 1412 VRYGFSGLG-GRDPGLSPGOELGRSVEELWALLSPILPGGALDRVLKXNLTAHSLDAQDS 1470
Db 1530 FRYGFSGLGSNSQALPPSHEVNDATIKQMKLLKLTQDTSADRFLSSGLGRFMAGLDTKNN 1589
QY 1471 LKTFWNNKGWHSVAVFNRRASNAILRAHLLPPGPARHAHSITTLNHPNLNTKEOLSEALM 1530
Db 1590 VKYWFNNKGWHAISFLNWINNAILRANLQKGENPSQYGITAFNHPNLNTKQOLSEVALM 1649
QY 1531 ASSVDVLVSTCVVVFAMSFVPASFTLVLIERVTRAKHLQMLGGLSPTLYWLGFLWDMCN 1590
Db 1650 TTSVDVLVSLCVFAMSFVPASFTVFLIQERSKAKHLQFISGVKPYIWLNFVWDMCN 1709
QY 1591 YLVPACIVVLIFLAFOQRAVYAPANIPALLLLLLLWGSITPLMYPASFPFSPSTAYV 1650
Db 1710 YVVPATLVIIIFCFOQKSVSSTNLPVLLALLLLWGSITPLMYPASFPFSPSTAYV 1769
QY 1651 LTCINLFIGINGSMAFVLELPSDQKLOEVSRLKOVFLIFPHFCIGRGLIDMVRNOAMA 1710
Db 1770 LTSVNLFIGINGSVAFVLELFTNNKLNINDILKSVFLIFPHFCIGRGLIDMVRNOAMA 1829
QY 1711 DAFERLGRDQFQSPFLXWVVGKLLAMVIOGPLFTLLQLHRSOLLPOPRVRSPLLG 1770
Db 1830 DALERGENRFPVSLNDVLGRNLFAMAVEGVVFTLVLIQYRFPFRPVKAKLPLN 1889
QY 1771 EDEDVARERVRVQATQGDVLVLRNLTKYRGQRMPADVRLCLGIPPGCEFLGVNG 1830
Db 1890 DEDEDVRRERQRIIDGGQNDILEIKELTKIYRRKRKPAVDRCIGIPPGCEFLGVNG 1949
QY 1831 AGKTSIFRWVTGTLARGEAVALAGHSVAREPAAHLSMGCYCPQSDAIPELLTGRHELE 1890
Db 1950 AGKSTTFKMLTGTPTVRGDAFLNKSILSNIEHVQNMGYCPQFDAITELLTGREHVEF 2009
QY 1891 LARLGVPEAQVTAQTAGSLARLGLSWYADRPAGTYSGGNKRKLATALALVGPVAVVFLD 1950
Db 2010 FALLRGVPEKEVGKFGWEAIRKGLVKYGEKYASNYSGGNKRKLSTAMALIGPPVAVVFLD 2069
QY 1951 EPTTGMDDPSARRFLWNSLLAVREGRSVMLTSHSMEECEALCSRLAIWNGRFRCLGSPQ 2010
Db 2070 EPTTGMDDPKARRFLWNCALLSIVKEGRSVLTSMSHECEALCTRMALWNGRFRCLGSPQ 2129
QY 2011 HLKGRFAAGHTLRLVDAARS--QPAAAFAVAEPGSELBEAHGGRRLRFOLPGGRCALA 2068
Db 2130 HLKNGRFDGTYIIVVIRAGSNPDLKPQVBFGLAPPGSVLKEKHKNMLQYOL-PSSLSLA 2188
QY 2069 RVFGEALVHCAEHGVEDFSVSTQMLBEVFLYFSKQCKDE---DTEQKEAGVGVDPA 2125
Db 2189 RIFESILSQSKRRLHIEDYSVSTQTLDOGVNFVFXKQSDDDHLKDLSHKQTV-VDVAV- 2246
QY 2126 LQHPKRVSQFLDDPSTAET 2144
Db 2247 -----LTSFLQDEKVKES 2259

RESULT 9
ABCI HUMAN
ID ABC1 HUMAN PRT; 2261 AA.
AC O95477; O96556; O96T85; O9NOV4; Q9UN06; Q9UN07; Q9UN08; Q9UN09;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette
DE transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux
DE regulatory protein).
GN Name=ABCA1; Synonyms=ABCL1, CERP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20345099; PubMed=10884428;
RA Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,
RA Freeman L.A., Cheng J.-F., Osorio J., Renaley A.T., Yang X.-P.,
RA Haudenschild C.C., Prades C., Chimini G., Blackmon E.E.,
RA Francois T.L., Duverger N., Rubin E.M., Rosier M., Deneffe P.,
RA Fredrickson D.S., Brewer H.B. Jr.;
RT "Complete genomic sequence of the human ABCA1 gene: analysis of the
RT human and mouse ATP-binding cassette A promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA "ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
RT regulated by LXR.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RA Qiu Y., Cavellier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies
RT revealing novel regulatory sequences.";
RL Genomics 73:66-76(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Tanaka A.R., Abe-Dohmae S., Arakawa R., Sadanami K., Kidera A.,
RA Kioka N., Amachi T., Yokoyama S., Ueda K.;
RT "A new topological model of functional human ABCA1-signal peptide
RT cleavage and glycosylation of a large extracellular domain.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99194549; PubMed=10092505; DOI=10.1006/bbrc.1999.0406;
RA Langmann T., Klucken J., Reil M., Liebisch G., Luciani M.F.,
RA Chimini G., Kaminski W.E., Schmitz G.;
RT "Molecular cloning of the human ATP-binding cassette transporter 1
RT (hABCL1): evidence for sterol-dependent regulation in macrophages.";
RL Biochem. Biophys. Res. Commun. 257:29-33(1999).
RN [6]
RP SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99364413; PubMed=10431238;
RA Rust S., Rosier M., Funke H., Real J., Amoura Z., Piette J.-C.,
RA Delucchi J.-F., Brewer H.B., Duverger N., Deneffe P., Assmann G.;
RT "Tangier disease is caused by mutations in the gene encoding ATP-
RT binding cassette transporter 1.";
RL Nat. Genet. 22:352-355(1999).
RN [7]
RP PHOSPHORYLATION SITES SER-1042 AND SER-2054.
RX MEDLINE=22289331; PubMed=12196520; DOI=10.1074/jbc.M204923200;
RA See R.H., Caday-Malcolm R.A., Singaraja R.R., Zhou S., Silverston A.,
RA Huber M.T., Moran J., James E.R., Janoo R., Savill J.M., Rigot V.,
RA Zhang L.H., Wang M., Chimini G., Wellington C.L., Tafuri S.R.,
RA Hayden M.R.;
RT "Protein kinase A site-specific phosphorylation regulates ATP-binding
RT cassette A1 (ABCA1)-mediated phospholipid efflux.";
RL J. Biol. Chem. 277:41835-41842(2002).
RN [8]
RP REPRESSION BY ZNF202.
RX MEDLINE=21192304; PubMed=11279031; DOI=10.1074/jbc.M100218200;
RA Forsch-Oezcuenermez M., Langmann T., Heimerl S., Borsukova H.,
RA Kaminski W.E., Drobnik W., Honer C., Schumacher C., Schmitz G.;

RT "The zinc finger protein 202 (ZNF202) is a transcriptional repressor
RT of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene
RT expression and a modulator of cellular lipid efflux.";
RL J. Biol. Chem. 276:12427-12433(2001).
RN [9]
RP INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not
RT ABCG1 via an LXR-independent pathway.";
RL J. Lipid Res. 43:952-959(2002).
RN [10]
RP REVIEW ON VARIANTS.
RX DOI=10.1161/01.ATV.0000078520.89539.77;
RA Singaraja R.R., Brunham L.R., Visscher H., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Efflux and atherosclerosis: the clinical and biochemical impact of
RT variations in the ABCA1 gene.";
RL Arterioscler. Thromb. Vasc. Biol. 23:1322-1332(2003).
RN [11]
RP VARIANTS HDL2 THR-1091 AND 1893-GLU-ASP-1894 DEL.
RX MEDLINE=20001430; PubMed=10533863;
RA Marcell M., Brooks-Wilson A., Clee S.M., Zhang L.-H., Yu L.,
RA Collins J.A., van Dam M., Molhuizen H.O.F., Loubser O.,
RA Ouellette B.F.F., Seneen C.W., Fichter K., Mott S., Denis M.,
RA Boucher B., Pimstone S., Genest J. Jr., Kastelein J.J.P., Hayden M.R.;
RT "Mutations in the ABC1 gene in familial HDL deficiency with defective
RT cholesterol efflux.";
RL Lancet 354:1341-1346(1999).
RN [12]
RP VARIANTS HDL1 ARG-597 AND ARG-1477, AND VARIANT HDL2 LEU-693 DEL.
RX MEDLINE=99364411; PubMed=10431236;
RA Brooks-Wilson A., Marcell M., Clee S.M., Zhang L.-H., Roomp K.,
RA van Dam M., Yu L., Brewer C., Collins J.A., Molhuizen H.O.F.,
RA Loubser O., Ouellette B.F.F., Fichter K., Ashbourne-Excoffon K.J.D.,
RA Seneen C.W., Scherer S., Mott S., Denis M., Martindale D.,
RA Frohlich J., Morgan K., Koop B., Pimstone S., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Mutations in ABC1 in Tangier disease and familial high-density
RT lipoprotein deficiency.";
RL Nat. Genet. 22:336-345(1999).
RN [13]
RP VARIANTS HDL1 SER-590; SER-935 AND VAL-937, AND VARIANTS ALA-399 AND
RP MET-883.
RX MEDLINE=99364412; PubMed=10431237;
RA Bodzioch M., Orso E., Klucken J., Langmann T., Boettcher A.,
RA Diederich W., Drobnik W., Barlage S., Buechler C.,
RA Porsch-Oezcuenermez M., Kaminski W.E., Hahmann H.W., Oette K.,
RA Rothe G., Aslanidis C., Lackner K.J., Schmitz G.;
RT "The gene encoding ATP-binding cassette transporter 1 is mutated in
RT Tangier disease.";
RL Nat. Genet. 22:347-351(1999).
RN [14]
RP VARIANTS HDL1 ILE-929; ARG-597 AND ARG-1477, AND VARIANTS HDL2
RP LEU-693 DEL; THR-1091; 1893-GLU-ASP-1894 DEL AND LEU-2150.
RX MEDLINE=20540002; PubMed=11086027;
RA Clee S.M., Kastelein J.J.P., van Dam M., Marcell M., Roomp K.,
RA Zwarts K.Y., Collins J.A., Roelants R., Tamasaawa N., Stult K.,
RA Suda T., Ceska R., Boucher B., Rondeau C., Desouich C.,
RA Brooks-Wilson A., Molhuizen H.O.F., Frohlich J., Genest J. Jr.,
RA Hayden M.R.;
RT "Age and residual cholesterol efflux affect HDL cholesterol levels and
RT coronary artery disease in ABCA1 heterozygotes.";
RL J. Clin. Invest. 106:1263-1270(2000).
RN [15]
RP VARIANTS HDL1 ASN-1289 AND HIS-1800.
RX MEDLINE=20171564; PubMed=10706591;
RA Brousseau M.E., Schaefer E.J., Dupuis J., Eustace B.,
RA Van Berdevegh P., Goldkamp A.L., Thurston L.M., FitzGerald M.G.,
RA Yasek-McKenna D., O'Neill G., Eberhart G.P., Weiffenbach B.,
RA Ordovas J.M., Freeman M.W., Brown R.H. Jr., Gu J.Z.;
RT "Novel mutations in the gene encoding ATP-binding cassette 1 in four

Db	1307	LLSGMDGKGSVQVGWKLTKQOFVALLWKRLIIARRSKGFPAQIVLPANPVCIALVPSL	1366
Qy	1259	IVPFPGHYPALRLSPTMYGAQVSPFSEDPAGDPGRARLLEALLQEAG-----LEE	1308
Db	1367	IVPFPGKYPSLQLOPMMYNEQYTVFSNDAPEDTGTLELLNALTKDPGFGTRCEBGNIPD	1426
Qy	1309	PPVQHSSHRFSAPEVPAEYAKVLASGNWTPESPSPACOCSQOPGARRLLPDCPAAAGPPP	1368
Db	1427	TPCQAGEEWTTAPVQPTIMDLFQNGWNWMTQNPSPACQCSDDKIKMLPCCPPGAGSLPP	1486
Qy	1369	PQAVTGSVEVQNLITGRNLSDFLVKTYPRVOGLKTKKWNEVRYGGSLSG-GRDPGLP	1427
Db	1487	PQRKQNTADIILODITGRNISDYLVKTYVQIIITAKSLKQKIWNNEFRYGGFSLCVSNTOALLP	1546
Qy	1428	SGQELGRSVEBELWALLSPGGLADRLVKNLITAWAHSLDQAQSLIKIWNFNKNGHSMVAFV	1487
Db	1547	PSQEVNDATQMKKHLKLAKDSSADRFLNSLGREMTGLDTRNNVKWNFNKNGHWAISFL	1606
Qy	1488	NRASNAIILRAHLPQGPABHAHSIITLHNHPLNLTKEOQLSEALMASSVDVLVSIQVAFMS	1547
Db	1607	NVINNAIILRANLQGENPNSHYGITAFNHPNLNLTKOQLSEVAPMTTSDVDLVISICVIFAMS	1666
Qy	1548	FVPASFVLITLIEERVTRAKHQLQMGSLPTLYLWGLNGLWDMCNLYLPACIYVLIFLAFQO	1607
Db	1667	FVPASFVVFLQEVRSKAKHLQFISGVKPYLYNLSNFWDMCNVYVPATLVIIIFICFQO	1726
Qy	1608	RAYVAPANLPALLLLLLLYGHSITPLMYPASFSPFSPSTAYVVLTCINLFIGINGSMATF	1667
Db	1727	KSYVSSNTNPLVALLLLLLYGHSITPLMYPASFVFEKIPSTAYVVLTSVNLFIGINGSVATP	1786
Qy	1668	VLELPSDOKLOEVSRIILKQVELIPPHCLGRGLTDMVNRQNMADAPERLGDROFQOSPLRW	1727
Db	1787	VLELPTDNKLNINILDKSVFLIPPHCLGRGLTDMVQNMADALERFGENFVPSPLW	1846
Qy	1728	EVWGKNLLAMVIOQSPFLFLLTLILQHSQSLLPQPRVRSPLPGLGEEDVARERERVVOGA	1787
Db	1847	DLVGRNLFAMAVEGVVFFLITVLIOYRFPTRPVNAKLSPLNDEDEDVRRERQRIIDGG	1906
Qy	1788	TQGDVLVLRLNTKYRGORMPAVDRLCLGIPPGCFGLLVNGVAGKTTSTFRWVTGDTLAS	1847
Db	1907	QONDILEIKELTKTYRRKRKPAVDRIICVGIIPGCFGLLVNGVAGKSTSTFKMLTGDTTVT	1966
Qy	1848	RGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVACTAG	1907
Db	1967	RGDAFLNKNSTLSNIEHVQNMGYCPOFDAITELLTGREHVEFFALLRGVPEKEVGKVE	2026
Qy	1908	SGLARLGLSWADRPAGTYSGGNKRKLATALLVGDPAVVPDDEPTTGMDBPSARRPLWNS	1967
Db	2027	WAIRKGLGVKGYEKYAGNYSGGNKRKLSTAWALLIGGPVVFLDEPTTGMDBPKARRFLWNC	2086
Qy	1968	LLAVVGRGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRPAAGHTLTLRVP	2027
Db	2087	ALSIVYKGRSVVLTSHSMECEALCTRMAIMVNGRFRCLGVSQHLKNRFGDGYTIVVRIA	2146
Qy	2028	AARS--OPAAAFVAABEPGSELREAHGRLRFQLPGGRCALARVFCGLAVHGAEGVED	2085
Db	2147	GSNPDLKPVQDFGLAPPGSVPEKHNMLQYQI--PSSLSSLAKRIFILSOSKKRLLHIED	2205
Qy	2086	FSVSQTMLEEVFLYPSKQDGKDE---DTEQKEAGVGDPAQGLQHPKRYVSQFLDDPSTA	2142
Db	2206	YSVSQTTLDQVFNFAKQSDDDHLKDLSLHKQTV--VDVAV-----LTSFLQDEKVK	2257
Qy	2143	ET 2144	
Db	2258	ES 2259	

RESULT 10
O80ZB2

Q00022 ID Q80ZB2 PRELIMINARY; PRT; 2201 AA.

AC Q80ZB2;
ID Q80ZB2;

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT	01-JUN-2003 (TREMBlRel. 24, Last sequence update)
21	01-JUN-2003 (TREMBlRel. 24, Last sequence update)


```
Qy 1864 AAHL5MGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAGTACGLARGLGSWYADRP 1923
Db 2005 DVH5MGYCPQSDAIFELLTGREHLYVARLGRVPADEIERVANWSIQSLGSLUYADRLV 2064
Qy 1924 GTYGGNKRKLATALALVGDPAVVLDEPTTGMDFPSARRFTLWNSLLAVVREGSRVMLTSH 1983
Db 2065 GTYGGNKRKLSTALALMGCPVLVLDDEPTTGMDFPSARRFTLWNSLLAVVREGSRVMLTSH 2124
Qy 1984 SMECEALCSRLA1MWNGRFCLGSPQHLKGRFAAGHTLTURVPAARS-----QPAAAF 2037
Db 2125 SMECEALCSRLA1MWNKTFQCLGTIOHLKYKFGDGYIVTMKFSKPKEDLLPDLNVPVEQF 2184
Qy 2038 VAAPFSGSEAREAGRLRFOLPGRCALARVGEALVGAHGAHVDEFSVQTMLEEVF 2097
Db 2185 FQGNFPGSVQRHYNMLQFQV---SSSLARIPLLLISHKDSLLIBEYSVTQTLDOQVF 2241
Qy 2098 LYFSKQDQKDRD 2109
Db 2242 VNAFAKQQTETHD 2253

RESULT 12
AAR87835
ID AAR87835 PRELIMINARY; PRT: 2269 AA.
AC AAR87835
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS23;
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY427778; AAR87835.1; -
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AS23;
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY427778; AAR87835.1; -
SQ SEQUENCE 2269 AA; 256510 MW; 28FE1DC0DE9E7BDB CRC64;

Query Match 45.1%; Score 5030; DB 2; Length 2269;
Best Local Similarity 45.6%; Pred. No. 9.7e-293;
Matches 1046; Conservative 358; Mismatches 666; Indels 222; Gaps 37;

Qy 1 MAPTQMLLWKNFMRRRQPVQVLLVPLFLPILVAVRSHHPLEHCHFPNKP 60
Db 1 MCFARQQLLWKNWTLRKQKIRFVVELVWPLSLFLILWLRNINFLYSQHECHFPNKA 60
Qy 61 LPSAGTVPLWGLCNVNNVTCFPLTPGEEPGRLSNFNDLSVSLLDARTVLGASAH 120
Db 61 MPSAGMLPWLQMCNVCNNPCFQNPPTGESGIVSNVNSILARVFRDFQELLIDAPERQ 120
Qy 121 TAGLGLKIATRAARSTAOPTKQSPLEPPMLDVAE----- 158
Db 121 HFHWKVEFQTLRLMDTLRTHPVRVAGRIRIRDLKDEETLFLMKNIGLSDSVVYL 180
Qy 159 LLSLLRTESL-----GLAQQAQ-- 177
Db 181 LLNSQVRPEQFAHGVPDMLMKDIACSETLLERFIIFSQRGAQTVRDAMCSLSQGLQWV 240
Qy 178 -----EPLSHLLEAAEDLAQELLARS-----LVELRALQRP----- 210
```

```
Db 241 EDTLIYANVDFPKLFRVLPTLLDSSSQGINLRSMGRVFSDISSRIREFIHRPSVEDLLWVT 300
Qy 211 ---RGTSGP-----LELLSEALCSV-RGFSSTVGSLSNMYEASDLMLVQGPESALP- 259
Db 301 KPLTQTGGPPTFAQLMSILSDLLCGYPEGGSRVF-SFNWYEDNNYKAFGLIDSTRKDPI 359
Qy 260 ---DSSLSPACSELIGALDSHPLSRLLKPLILGLKLLFAPDTPPTRKLMQAVNTEFE 316
Db 360 YSDVKRTTTCNALIQSLESNPITKIARAAKPLVMGKILFTPDSPAVRILQVANTEFE 419
Qy 317 ELTLRLDRVREVMELGPRIETFMNDSNVAMLRLL-----QMODEGRQRPRPG 365
Db 420 ELERLRLKLVKAEVGGQIWFYFDRSTQMTMIRDTLENPTVKGFNLSQLCEGIT----- 474
Qy 366 GRDHMEALRFLDPG-----SGYSWQDAHADVGHVLTGRVTECLSLDKLEAAPS 417
Db 475 ---AEAMLNFKGPRESQADDMANFDWRDVFNIITDRTLRLTSKYLECILDKFESYDD 530
Qy 418 EAALVSRALOLLAHREFWAGVWVFLGPDSDPTHEPTDGLPGHVRIRKIMDIDVVRTN 477
Db 531 EIQLTQALSLLEENRFWAGVVF-----PDNYFWTSALPTHVKYKIRMDIDVVEKTN 582
Qy 478 KIRDREWDPCGAADPLTDLRYVMGGFYVLQDLVERAAVRVLSGANPRAGLYLOMPYCY 537
Db 583 KIRDYWDSCPRADPVEDFRYIWGGPAYLQDMIEQGITRSQAQVEVPVGIYLOMPYPCF 642
Qy 538 VDDVFLRVLRSILPLFLTLAWIYSVILTVKAVREKETRLDRMTWRANGLSRAVLWGLWFL 597
Db 643 VDDSFMIILNRCFPIFMVLAWIYSVMTVKSIVLEKELRLKETLKNQGVSNVAVTCWTFWL 702
Qy 598 SCLGPFLLSAAALVVLKGLDILPYSHPGVVFLLAFAFATVATVTSFLLSAFTSRANLAA 657
Db 703 DSFSIMSMSIFLITIFIMHGRILHYSNPFILFLFLAFSTATTMOQCLLSTFTFSRASLAA 762
Qy 658 ACGGLAYFSLYLPVLCVAMRDLRPGGRVAAALLSPVAFGFCGESLALLEEQEGEAQWH 717
Db 763 ACSGVITYFLYLPHILCFAMQDRMTADLKMAVSLSPVAFGFGTEYLARFEEQGLQWS 822
Qy 718 NVGTRP-TADVLSLAQVSGLLLDAAALYGLATWYLEAVCGQYIGIPPEWNPFRSRWCG 776
Db 823 NIGKSPMEGDEFSFLSMKMLLDAAALYGLLAWLDQVFPNGTGTPLPWFLLQESYWL 882
Qy 777 -----PRPKSPAPCPTPL-DPK-----VLVEEAPPGLSPGVSRVSRLEKRFPGSP 820
Db 883 GEGCSTREERALEKTEPITEEMEDDPHEPEGINDAFFERELPGLVPGVCVKNLVKIFPYS 942
Qy 821 QPALRGLSLDFYQCHITAFIAGHNAGKTTTSLISGLFPPSGGSFAFLGHNDVRSSMAAIR 880
Db 943 RPAVDRLNITFYENQITAFIAGHNAGKTTTSLISGLFPPSGGSFAFLGHNDVRSSMAAIR 1002
Qy 881 PHLGVCPOYNVLFMDLTVDEHVVWFYGLKGLSAAVVGPEQDRLLLODVGLYSKQSVQTRHL 940
Db 1003 QSLGMCPOYNILFHLTVAEHILFYAQLKGKSEEAQLEAMEAMLEDTGLHKKRNEEAQDL 1062
Qy 941 SGGWOKRLSVAIAPVGSQVVIDEPTAGVDPASRRGIWELLKRYRGRITLILSTHLE 1000
Db 1063 SGGWOKRLSVAIAPVGSQVVIDEPTAGVDPASRRGIWELLKRYRGRITLILSTHLE 1122
Qy 1001 AELLGRVAVVAGRLCCCGSPFLRRHLGSGYLLTLVKARLPLTTNEKADTDMEGSVDT 1060
Db 1123 ADLLGRIALISQRLVCSGTPLFKNCFGTGFTVLVR---KMKNIQSQTGCEGTCSC 1179
Qy 1061 ROEKNGSQSGRVG--TP-----QILLALVQHVWVPGARIVBELPHELVLVLPYTGAAH 1109
Db 1180 ASEGFSTRCPAHIDEITPEQVLDGVDNMLDMVQHVVPEAKLVEICQGLIFILLPNKNFK 1239
Qy 1110 DGSFATLRFELDTFLAELRUTGYISDTSLEEFKLVVERCAADTDMEDSCGHLCTGI 1169
Db 1240 QRAYASLFRLEEDTLADLGLSSFGISDTPLLEEFLKVT-----EDSDSGPLFAGGT 1290
Qy 1170 AGLDVTLRK-MPPQETALE-----NG---EPAGSAPETDQSGPDVAGRVQ-GWALTRO 1219
Db 1291 QOKRENLRHPWLPSPREKARQIPQSGNGCSXEPA-PHPEGQPSSEPEARSLRINTGAQLIIQ 1349
```


Db 531 EIQUTQALSLLEENRFWAGVVF-----PDMYPWTSALPTVKKYKIRMDIDVVEKTN 582
Qy 478 KIRDFWDGPAADPLTDLRYVWGGFVYQLDVRRAAVRVLSGANPRAGLYLOMPYPCY 537
Db 583 KIKDRYWDGPRADVEDFEDRYINGGFAYLQDMI BQGITRISOAQVKVPVGIYLOMPYPCF 642
Qy 538 VDDVFLRVLRSPLPLFLTLAWIYVTLTVKAVVREKETRLDRWRANGLSRAVLWGLWFL 597
Db 643 VDSFMIILNRCFFIFMVLAWIYVSMTVKSIVLEKELRLKETLKNQGVSNVITWCTWFL 702
Qy 598 SCLGFPLLSAALLVVLKGLDILPYSHVGVVFLFAAFVAVATVQOSLLGAFFGRANLAA 657
Db 703 DSPSIMGNSIFLLTIFIMHGRILYHNSPFIULFLLAFSTATINQCFLLSTFFSASLAA 762
Qy 658 ACGLAYFSLYPLVYLVCLVNRDLRPAAGRVAASLLSPVAFGFCESLALLEEQEGEAQWH 717
Db 763 ACSGVYFTLYLPHILCFAMQDRMTADLKMAVSLSPVAFGFTGYLARFEEQGLQWS 822
Qy 718 NVGTRP-TADVFLSAQVSGLLLLDAAALYGLATWYLEAVCPQOYCIPEPWPFPRRSWCG 776
Db 823 NIGKSPMEGDFSLMSKMWMLLDAALYGLLAWLDQVFPENYGTPLFWYFLLOESYWL 882
Qy 777 -----PRPKSPAPCPTPL-DPK-----VLVEAPPCLSPGVSVRSLEKRFPGSP 820
Db 883 GEGCSTREAREALEXTEPITEEMEDPEHPEGINDAFFERELPCLVPGVCVKNLVKIFEPYS 942
Qy 821 OPALRGSLDFYOGHITAFIIGHNGAGKTTTILSILSGLFPSPGSAFILGHDRSSMAAIR 880
Db 943 RPAVRLNITFYENQITAFIIGHNGAGKTTTILSILSGLFPSPGSAFILGHDRSSMAAIR 1002
Qy 881 PHLGVCQYVNLFDMLTWDHWHVYGRILKGLSAAVGPEDQRLLODVLGVSQSVQTRHL 940
Db 1003 QSLGMCQYVNLFDMLTWDHWHVYGRILKGLSAAVGPEDQRLLODVLGVSQSVQTRHL 1062
Qy 941 SGGMORKLSVAIAFVGGSOVVIDEPTAGVDPASRRGIWELLKYRREGRTLILSTHLE 1000
Db 1063 SGGMORKLSVAIAFVGGAKVVIDEPTAGVDPASRRGIWELLKYRREGRTLILSTHLE 1122
Qy 1001 AELLGDRVAVVAGRLCCGSPFLRRHLGSGYLYTLVKARLPLTTNEKADTDMEGSVDT 1060
Db 1123 ADLLGDRIAIISQGRLYCSGTFPLKNCFGTGYTLVR---KMKNIQSQTGCEGTSC 1179
Qy 1061 ROEKNKSGSRVG--TP-----OLLALVGHVWPGARLVEELPHELVLVLPYTGAAH 1109
Db 1180 ASEGFSTRCPAHIDEITPEQVLDGVNLMQVHHVPEAKLVCEICQELIFLLPNKFX 1239
Qy 1110 DGSFATLFRLEDRLAELRLTGYSIDTSLSEIEFLKVVEECAADTMDGSCGQHLCTGI 1169
Db 1240 QRAYASLFRLEDRLADLGLSSFGISDTPLEIEFLKVT-----EDSDSGPLFAGGT 1290
Qy 1170 AGLDVTLRLK-MPQETALE-----NG---EPAGSAPETDGGSPDAVGRVQ-GWALTRQ 1219
Db 1291 QOKRENLRHPWLSPREKARQIPQGSNGSCXEPA-PHEPGQSSPEARSRLNTCAQLIIQ 1349
Qy 1220 QLQALLKRFLLARRSRGLFPAQIVLPAFVGLALVSLIVPPCHYPALRLSLTMYCAQ 1279
Db 1350 HVQALLKRFHHTIRSHKDFLAQIVLPAFVFLALMLSIIVPPCEYPAALILHPMYGQ 1409
Qy 1280 VSFFSEDAAGPPGRARLLEALLQAG-----LEEPPVOHSHRFSAPVPAVAK 1329
Db 1410 YTFSLQPGSEQLAALAVLLNKPFGNRCLEKLEWLPPEYPCGNST-PWKTSPSVNITH 1468
Qy 1330 VLASGNWTPSPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTSGSEVQNLTRNLSD 1389
Db 1469 LFQKQWTPKPSPCRSCTREKLTMLPECEGAGGLPPQRIQRSTEILLQDLTNRNID 1528
Qy 1390 FLVITYRLVRQGLTKTKKWNVEVYGGFSLGCRDPGLP-SQOELGRSVEELWALLSPLP 1448
Db 1529 FLVKTYPALTRSSLSKSFVWNEQYGGISIGGKLPILPITGEALVEFLSHLQIMN-VSG 1587
Qy 1449 GALDR-VLKNLTAWAHSLDAQDSLKIFWNNKGHSMVAVFVNRASNAFLRAHLPPGPARHA 1507

Db 1588 GPITREASKEMPAPFLKHLETEDNIKVWFNNKGHWALVSLFVNAHNAIURTSLHKDKNPEE 1647
Qy 1508 HSTITLNLPLNLTKEQLSEALMASSVDVLVSVICVFPAMSFVPASFTLVILIEERVTRAKH 1567
Db 1648 YGITVISQPLNLTKEQLSEITVLTASVDAVAICVIFPAMSFVPASFTLVILIEERVTRAKH 1707
Qy 1568 LQJMGGLSPTLYWLGHNFLNMDCNVLPACIVLILFLAQORVAVAPANLALLLLLYG 1627
Db 1708 LQFVSGSPSTYTLWTLNFDIMYAYSAALVGVFGFKKAYTSPENLPALIALLMY 1767
Qy 1628 WSITPLMYPASFFSPSTAYVVLTCINLFIGINGSNATFVLEFSDQ-KLQEVSRILKQ 1686
Db 1768 WAVIPMYPASFLFDVPSTAYVALSCANLFIGINSSAITFILELFENNWTLLRRENAMLRK 1827
Qy 1687 VFLIFPHFCLGRGLIDMVRNOAMADAFERLGDQFQFQSPLRWEVVGKLLAMVQGPFL 1746
Db 1828 LLIIIFPHFCLGRGLIDALASQAVTDVYARFGEHSTNPFQMDLIGKNLVAMAAEGVVYLL 1887
Qy 1747 FTLLQHR---SOLLPOPRVRSPLLGEDEEDVARERERVQCATGCDVLVLRNLTKVYR 1803
Db 1888 LTLFIQHFFLRTWSEPAKE--PII-DEDDVAERQRIISGKNKTDILRLNELTKIYS 1944
Qy 1804 GQMPAVDRICLGIIPGCECFGLGVNGAGKSTFRMVGTDLASRGEAVLAGHSVAREPS 1863
Db 1945 GTSSPAVDRICVVRPCECFGLGVNGAGKSTFRMVGTDLASRGEAVLAGHSVAREPS 2004
Qy 1864 AAILSWGYCQSDAIFELLTGRHEHLELRLARLGRVPEAQVATAGSLARLGLSWYADRP 1923
Db 2005 DVHQSMGYCQSDAIFELLTGRHEHLELRLARLGRVPEAQVATAGSLARLGLSWYADRP 2064
Qy 1924 GTYSGGNKRLATALVAGDPAVVVFLDEPTTGMDPSARRFLNWSLLAVVREGSRVMTSH 1983
Db 2065 GTYSGGNKRLATALVAGDPAVVVFLDEPTTGMDPSARRFLNWSLLAVVREGSRVMTSH 2124
Qy 1984 SMEECEALCRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLVPAARS-----QPAAAF 2037
Db 2125 SMEECEALCRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLVPAARS-----QPAAAF 2184
Qy 2038 VAAEFGSELRHAGRLRQLPGRCALARVFGELAVHGAHGVDFSVQTMLEVF 2097
Db 2185 FQGNFPGSVQRHRYNMQFV---SSSLARIFRLISHKDSLLIEBYSVTQTLQDF 2241
Qy 2098 LFESKQDQKDED 2109
Db 2242 VNFAKQOTETHD 2253

RESULT 14
AAR87836 PRELIMINARY; PRT; 2269 AA.
AC AAR87836; 20-MAY-2004 (TReMBLrel. 27, Created)
DT 20-MAY-2004 (TReMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.;
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.";
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P1117;
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY427779; AAR87836.1; -

SQ SEQUENCE 2269 AA; 256491 MW; 956E349FF17895A1 CRC64;

Query Match

Best Local Similarity 45.1%; Score 5029; DB 2; Length 2269;

Matches 1046; Conservative 358; Mismatches 666; Indels 222; Gaps 37;

QY 1 MAFPTQLMLLWKNFYRRQPVQLLVLLVLLFFLFFLVAVRHSHPPLEHECHFPNKP 60
DB 1 MGFAQIQILLWKNWTLRKQIRFVVELVWPLSLFLILLIWRNINPLYSQHECHFPNKA 60
QY 61 LPSAGTVPWLOGLICNNVNTCPQLTPGEBPGRLSNFNDLSVRLADARTVGGASARR 120
DB 61 MFSAGMLPWLOGMFCNNVNPFCQNPPTPGSPGIVSNYNNISILARFRDFOELLDAPEQ 120
QY 121 TLAGLGKLTATLRAARSTAPQPTKQSPLEPPMLDVAE----- 158
DB 121 HFGHWKQFQTLRLMDTLRTHPVRVAGRGIRDLVLDKEETLTLFLMKNIGLSVSVYL 180
QY 159 LLTSLRLTESL-----GLALGOA-- 177
DB 181 LLNSQVRPQPAHGVDPDLMLKDIACTETLLERFIIFSQRGAQTVRDAMCSLSQGTLOW 240
QY 178 -----EPLHSLLEAAEDLAQELLALRS-----LVLEALLQRP----- 210
DB 241 EDTLVANVDFFKLFRVLPDLLDSSSQGINLRSGWRVFSDISSRIREFIHRPSEVDLLWT 300
QY 211 ---RGTSGP-----LELLSEALCSV-RGPSVTGSPSLNMYEASDLMELVQEPESALP- 259
DB 301 KPLTQTGGPBTFAQLMSILSDLCGYPEGGSGRVF-SFNWYEDNNYKAFGLGIDSTRKQPI 359
QY 260 ---DSLSLSPACSELGALDHPSLRLWRRLKPLILGKLLFAPDPTFKLMAQVNRTPF 316
DB 360 YSYDKRTTFNCALQSLNESPLTKIARPAAPLWVGKILFTPDSPAVRRIILQANSTFE 419
QY 317 ETLTLDVREVMELGPRIFTFMDNSSNVAMLQRL-----QMDREGRRQPRPG 365
DB 420 ELERLKLKWAEEVGQPIWYFDRSTQMTMIRDITLENPTVKGFLNSQLGEBGII----- 474
QY 366 GRDHMEALRSFLDPG-----SGYSWODAHADVGLVGTGRVTECLSLDKLEAAPS 417
DB 475 ----AEAMLNFLHKGPRESDQADMANFDWRDVFNTDRTLRLTSKYLECLILDKFESYDD 530
QY 418 EALVSRALQLLAEHRFWAGVFLGPEDSSDTEHTPTDLGPHGVRIKIRMDIDVVRTN 477
DB 531 ELQTLQALSLLEERFWAGVVF-----PDMYPTWSALTPTHVKYKIRMDIDVVEKIN 582
QY 478 KIRDFWDGPAADPLTDLRYVWGFVYLQDLIVERAAVRLVSGANPRAGLYLQMPYPCY 537
DB 583 KIKDRYWDGSPRADPVEDPRYTWGGFAVLQDMIEQGITRQAQVXVPVGVYILQMPYPCF 642
QY 538 VDDVFLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLRDTMRAMGLSRAVLWGLWFL 597
DB 643 VDDSFMIILNRCFPIFMWLAWIYSVMTKSVLEKELRLKETLKNQGVSNVAVICTWFL 702
QY 598 SCLGPFPLSALLVLVLKGLDILPYSHPGVWFLAAPAVATVTOSEFLLSAFESRANLAA 657
DB 703 DSFSIMSMISFILTTFIMHGRILHYSNPFILFLFLLAFSTATIMQCFLSTFFSRASLAA 762
QY 658 ACGGLAYFSLYPLVLCVAVMRDLRPPAGGRVAAASLSPVAFGFCESLALLERQSGAOWH 717
DB 763 ACSGVITYFTLYLPHILCFPAQWDRMTADLKMAVSLSPVAFGFGTEVLAARFEEQGLQWS 822
QY 718 NVGTTP-TADVPSLAQVSGLLLDLDAALYGLATWYLEAVCPGQGYPEBWNPFPRSYWCG 776
DB 823 NIGKSPMEGDEFSFLMSMKOMMLDAALYGLLAWYLDQVPPGNYGTPLPWYFLLOSXYWL 882
QY 777 -----PRPPSPAPCTPL-DPK-----VLVEBAPPGLSPGVSRLEKEFPGPS 820
DB 883 GEGCSTREERALEKTEPTTEMEDPEHPGINDAFPERELPGLVPGVCVKNLVLKLPSPYS 942
QY 821 QPALRGLSLDFYQGHITAPLGHNGAGKTTTTLTSLSGLPFPPSGSAFILGHVDRSSMAAIR 880

DB 943 RPAVDRLNITFYENQITAPLGHNGAGKTTTTLTSLGTPPTSGTVLIGKDIETSLDAVR 1002
QY 881 PHLGCPQYNVLPFDMLTVDYEHVWFYGRUKGLSAAVVGPEQDRLLODVLGVLSQSVQTRHL 940
DB 1003 OSLGMCPOQNLILFHHLLTVAEHLFLYAOLKGSWEBAQLEAMEALDEDTGLHHRKNEBAQDL 1062
QY 941 SGMQRKLSVAIAFVGGSQVWILDEPTAGVDPASRRGIWELLKLYKREGRTLLSTLTHLDE 1000
DB 1063 SGMQRKLSVAIAFVGGAKVILDEPTSGVDPSRRSIWDLKLYKRSRTIIMSTHMD 1122
QY 1001 AELLGDRVAVAGGRCCGSPFLFRRHLLGSGYYITLVKARLPLTTNEKADTDMGSDVT 1060
DB 1123 ADLLGDRITAIISQGRLYCGTFLKNCFCGTGYTLTVR---KMKNIQSORTGCEGTSC 1179
QY 1061 ROEKNGSQSGRVG--TP-----OLLALVOHWVPCARLVEBELPHELVLVLPYGAH 1109
DB 1180 ASEGSTRCPAHIDITPEQVLDGVDNEMDMVQHVPEAKLVCEIGQELIFLKNKPK 1239
QY 1110 DGSFATLPRELDTRLAELRLTGYGISDTSLSBEIFLVVVECAADTDMEDGSCGQHLCTGI 1169
DB 1240 QRAYASLFRELEDTLADLGLSPGISDTPLEIFLKV-----EDSDGSLFAGGT 1290
QY 1170 AGLDVTLRK-MPPOETALE-----NG---EPAGSAPETDQSGSDPAVGRVO-GWALTQ 1219
DB 1291 QOKRENLRHPWLSPREKARQIPQSSNGCSXSPA-PHPEGQPSSEPEARSRLMTGAQLITQ 1349
QY 1220 QLOALLKXRFLLARRSRGLFAQIVLPALFVGLALVFSIIVDPGHYPALRLSPMYGQ 1279
DB 1350 HVQALLVRRFHTTRSHKDFLAQIVLPATVFLPALMSIIVPPFGEYPALIILHPMYGQ 1409
QY 1280 VSFFSEADPGDGRARILLEALLQEAQ-----LEPPVQHSRSHPSADEVPAEVAK 1329
DB 1410 YTFSLDQPSQEAALADVLNKGFGNRCLEKEMLPEYPCNST-PKWTSPSPNITH 1468
QY 1330 VLASGNWTPSPSPACQSCQFARLLPDCPAAAGPPPPQAVTSGSEVQVQLTGRNLS 1389
DB 1469 LFOQWTPKPSPCRSCTREKLTMLPECPEGAGLPQKRIQRTSILQDLTNRNID 1528
QY 1390 FLVKTYPPLVROGLTKMNVNVRVGGSLGDRDPLP-SQELGRSEVELWALLSLPLG 1448
DB 1529 FLVKTYPALIRSSLSKSFVNEQRYGSIIGKLPILFITGALVEFLSHLQIWN-VSG 1587
QY 1449 GALDR-VLKNLTAWAHSILDAQDSKIFWNNKGMWMAFVNRASNAIILRAHLPPGPARHA 1507
DB 1588 GPIREASKEWPAFLKLETDENIKVFNKGMWMAFVNRASNAIILRAHLPPGPARHA 1647
QY 1508 HSITTLNHLNLTKEQLESAALMASSVDVLVSIQVFPAMSPVPSFTLVLBEERVTRAKH 1567
DB 1648 YGITVISQPLNLTKEQLESEITVLTASVAVAIQVIFAMSPVPSFVLVLIQERNVAKH 1707
QY 1568 LQMGSLSPTYLWGLNFMNLYVPACIVVLIFLAFQRAYAPANIPALLLLLLLYG 1627
DB 1708 LQFVSGVSPTYLWGLNFMNLYVPACIVVLIFLAFQRAYAPANIPALLLLLLLYG 1767
QY 1628 WSITPLMYPASFFSPSTAYVLTCLNLTGINGSMATFVLELPDQ-KLOEVSRLKQ 1686
DB 1768 WAVIPMYPASFFSPSTAYVLTCLNLTGINGSMATFVLELPDQ-KLOEVSRLKQ 1827
QY 1687 VFLIPPHCLGRGLIDMVNRQAMADAFERLGRDQSPSPRLWEVVKNLAMVYIQSPFL 1746
DB 1828 LLIIPPHCLGRGLIDLALSAQVTDVYARFGEBSHTNPFQWDLIGKLVMAAEGVYLL 1887
QY 1747 FTLLLOHR---SOLLQPVRSLSPLGGEEDVABERERVQGTQGVVLRLNLTQYR 1803
DB 1888 LTLFQHHFFLTRWVSEPAKS--PII-DEDDVABERQRIISGGNKTDIRLNEUTKYS 1944
QY 1804 GORPAPVRLCLGIPPGECFGLGVNGAGKTTSTFMVTGDTLASRGEAVLAGHSVARBPS 1863
DB 1945 GTSSPAVDRLCVGRPGECFGLGVNGAGKTTTFKMLTGDITVTSGDATLAKSLITNIS 2004
QY 1864 AAHLSMGYCPQSDAIFELLTGRHLELRLGRVPEAQVTAQTAGSGLARLGLSWYADRP 1923

Db 2005 DVHQSNGYCPQFQDAVDDLLTGREHLYLYARLRGVPADIEIRVANKWSIQSLGLSLYADRLV 2064
Qy 1924 QTYSGNKRKLATLALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGSVMLTSH 1983
Db 2065 QTYSGNKRKLSTALMGCPPLVLLDPTTGMDPQARRMLWNTIVSIREGRAVLTSH 2124
Qy 1984 SMEECALCSRLATMNGRFRCLSGPOHLKGRFAAGHTLTLRVPAARS-----QPAAAF 2037
Db 2125 SMEECALCTLRATMVGRTFQCLGTIQLKTKYFGDGYIVTKMKSPKEDLLPLDNPVEQF 2184
Qy 2038 VAAFPFSGELREAHGRLRFQPPGRCALARVFGELAVHGAEGVEDFSVSTQMLBEVF 2097
Db 2185 FQGNFPGSVQERHYNMLOFV---SSSSLARIFFLLISHKDSLLIBEYSVTQTLDOVF 2241
Qy 2098 LYFSKQDQKDED 2109
Db 2242 VNAFAQQTETHD 2253

RESULT 15
Q6T942
ID Q6T942 PRELIMINARY; PRT; 2269 AA.
AC Q6T942;
DT 05-JUL-2004 (TreeMRel. 27, Created)
DT 05-JUL-2004 (TreeMRel. 27, Last sequence update)
DT 05-JUL-2004 (TreeMRel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064680;
RA Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
RA Aguirre G.D., Acland G.M.;
RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
RT dystrophies and progressive retinal atrophies.";
RL Mol. Vision 10:223-232(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
RA Acland G.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY427777; AAR87834.1; -.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005951; Rim_ABC_transpt.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR01257; rim_protein; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2269 AA; 256482 MW; 00AF870E6D9BE9A5 CRC64;

Query Match 45.1%; Score 5028; DB 2; Length 2269;
Best Local Similarity 45.6%; Pred. No. 1.3e-292;
Matches 1046; Conservative 357; Mismatches 667; Indels 222; Gaps 37;
Qy 1 NAFWTQLMLLWKNPMYRRRQPVQLLVLLPLFLFFILVAVRSHHPLEHCHFPNKP 60
Db 1 MGFARQIQLLWKNWTLRKQKIRFVVELVWPLSLFLILWLRLNINPLYSQHECHFNKA 60
Qy 61 LPSAGTVPWLOGLICNVANTCPQLTGCEEPGRLSNFENDSLVSLLDADARTVLGASAH 120
Db 61 MPSAGMLPWLOGMFCNVNPNCPQNTPGESPGIVSNYNNSILARVFRDFQELLADAPERQ 120
Qy 121 TLAGLGLKIATLRAARSTAQOPTKQSPLEPPMLDVAE----- 158
Db 121 HFGHWKKEFQTLRLSLMDTLRTHPervAGRGIRIRDLKDEETLTTLFLMKNLGLSDSVVYL 180

Qy 159 LITSLLRATESL-----GLALGOAQ-- 177
Db 181 LINSQVRPEQFAHGVDPDLMDKIACSTLLERFIIFSQRGAQTVRDAMCSLSGTLQWV 240
Qy 178 -----EPLHSLLEAAEDLAQELLALRS-----LVELRALLQRP- 210
Db 241 EDTLYANVDFPKLFRVLPTLLDSSSQGINLRSWGRVFSDISSRIREFIHRPSVEDLLWVT 300
Qy 211 ----RGTSGP-----DELLSEALCSV--RGPSSVTGPSLNWYEASDLMELVQGPESALP- 259
Db 301 KPLTQTGCTPETAQLMSILSDLLCGYPEGGSRVF--SFNWEYDNNYKAFGLIDSTRKDPI 359
Qy 260 ----SSSLSPACSELI GALDSHPLSRLLRLLKPLILGLKLPADPTTRKLMQAVNTEFE 316
Db 360 YSYDKRTTTFNCALIQSLESNPLTKIAWRAAKPLVMGKILFTPDSPAVRILQANSTFE 419
Qy 317 ELTLRLDRVREVMELGPRIFTFMNDSNVAMLOQLL-----QMODEGRRQPRPG 365
Db 420 ELERLKLKVANEVGPQIWIYFDRSTQMTWIRDTLENPTVKGFLNSQLGSEGIT----- 474
Qy 366 GRDHMEALRSFLDPG-----SGGYSWQDAHADVGHLYGTILGRVTTECLSLDKLEAAPS 417
Db 475 ----AEAMLNFLHKGPPRESQADDMANFDWRDVFENITDRTLRLTSKYLECLILDKFESYDD 530
Qy 418 EAAVLSPALOLLAEHREWAGVFLGPDSDPTEHPTPDLPGLGHVRIKIRMDIDVVRTN 477
Db 531 EIQTQTSALSLLEENREWAGVVF-----PDMYPMWTSALPHTHVKYKIRMDIDVBEKTN 582
Qy 478 KIRDRFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPY 537
Db 583 KIKDRYWDSPRADPVEDFRIYWGGFAYLQDMLEQGITRSQAQVEVPVGIYLOMPYPCF 642
Qy 538 VDDVFLVLSRLPLFLTLLAWIYSVTITVKAIVREKETRLDRTRANGLSRAVLWGLWFL 597
Db 643 VDDSFMIILNRCFPFIMVLAWIYSVMTKSI VLEKELRLKETLKNQGVSNVITWCTWFL 702
Qy 598 SCGLGPFLLSALLVLVLKGLDILPYSHPGVVFLLAFAFATVYTSQSLLSAFSTRANLAA 657
Db 703 DSFSIMSMSTIFLTFIMHGRILHYSNPFILFLFLAFSTATIMQCFLSTFFSRASLAA 762
Qy 658 ACGLAYFLSLPYLVLCVAVNRDLRPGAGRVAAASLLSPVAFGFCESIALLEEQEGEAQWH 717
Db 763 ACSGVITYTLYLPHILCFAMQDRMTADLKAVSLLSPVAFGFGTEYLARFEEQGLGLQWS 822
Qy 718 NVGTRP--TADVSLAQVSGLLLDAAALYGLATWYLEAVCPQGYGIPFPWNPFRSRWCG 776
Db 823 NIGKSPMEGDEFSPLMSKMWMLLDAALYGLLAWYLDQVFPQNYGTPLPFWYFLLOESYWL 882
Qy 777 -----PRPKSPACPTPL-DPK-----VLVEEAPGLSPGYSVRSLEKRFPGSP 820
Db 883 GEGCSTREERALEKTEPITEEMEDPEHPEGINDAFFERELPGLVPGVCVKNLVKIFBPYS 942
Qy 821 QPALRGLSLDFYOGHITAFILCHNGAGKTTLSILSGLPPSPGSAFSLGHGHDVRSMAAIR 880
Db 943 RPAVDRLNITFYENQITAFILCHNGAGKTTLSILTGLPTSGIVLIGKDIETSLDAVR 1002
Qy 881 PHLGVCQYNVLFDMLTVDHEWVWFYGRKLGLSAAVVGPQDRLLQDVGLSVKQSVQTRHL 940
Db 1003 QSLGMCQYNNILFHLHTVAEHILFYAQLKGKSWEEAQLMEAMLEDGLAHKRNEEAQDL 1062
Qy 941 SGGHQRKLSVAIAFVGSQVVIDDEPTAGVDPASRRGIIWELLKRYRSGRTILSTHLD 1000
Db 1063 SGGHQRKLSVAIAFVGSQVVIDDEPTAGVDPASRRGIIWELLKRYRSGRTILSTHLD 1122
Qy 1001 AELLGDRVAVVAGRLCCGSPFLRRHLGSGYVTLVKARLPLTTNEKADTDMEGSVDT 1060
Db 1123 ADLLGDRALIASQRLYCSGTPLFKNCFGTGFVTLVR---KMKNTQSQTGCEGTCSC 1179
Qy 1061 ROEKNGSQSGSRVG--TP-----QLLALVOHVWVPGARLVEELPHELVLVLPYGAH 1109
Db 1180 ASEGFSTRCPAHIDEITPEQVLDGVNLMQVHVPVPEAKLVEICQOELIFLLPNKNFK 1239

Qy	1110	DGSFATLPRFLDTRLAELRLDTCYGISDTSLEEEIPLKVVEECAADTDMDSCSGOHLCTGI	1167
Db	1240	QRAVAFPRELDTLADLGLSSFGSDTFLPEEIFLKVT-----EDSDGSLFAGGT	1290
Qy	1170	AGLDVTLRLK-MPPQETALE-----NG-----EPAGSAPETDQSGDPDAVGRVO-CWALTRO	1219
Db	1291	QOKRENLRHPWLSPREKARQIPQSGNGCSAEPA-PHPEQGQPSSEPEARSRLNTGAQLIIQ	1349
Qy	1220	QLOALLKRFLLARRSRGLPAQIVLPALFVGLALVFLSLIVPPPGHYPAIRLSTMYGAQ	1279
Db	1350	HVQALLVKRFHTTIRSHKDFLQAIQVLPATFVFLALMSIIVPPPEGYFALLIHPMYGQQ	1409
Qy	1280	VSPFSEADGPDGGRARLEALLQBAQ-----LEEPVQHSRPSAEPEVBARVAK	1329
Db	1410	YTFPSLQPGSEQLAALADVLNKPFGNRCLEKWLPEYPCGNST-PWKTVSPSPNITH	1468
Qy	1330	VLASGNWTPSPSPACQCSQPGARRLLPDCPAAAGGPPPPQAVTGSGEVVQNLTRNLSD	1389
Db	1469	LFQKQWOTPEKPSPCSCSTREKLTMLPECEGAGGLPPQRIQRSTEILQDLTRNRISD	1528
Qy	1390	FLVKTYPRLVROGLKTKWNEVRYGFSLGGRDGGLP-SGOELGRSVEELWALLSLPLG	1448
Db	1529	FLVKTYPALIRKSLSKSPFWNEQRYGSGIGGKLPILPITGEALVEFLSHLGGIMN-VSG	1587
Qy	1449	GALDR-VLKNLTAWAHSIDAQDSLKIWFNNKGHSMVAFVNRASNATLRAHLPPGPARHA	1507
Db	1588	GPIITREASKEMPAFLKHELTEDNIKVFNNKGWHALVSLNVAHNALTRLSLHKDKNPEE	1647
Qy	1508	HSITLNLHPLMTKEQISEAALMASSVDVLVSCVFPAMSFVPASFTVLVIERVTRAKH	1567
Db	1648	YGITVISQPLNLTKEQISEITVLTAQSDVAIVAICVIPAMSFVPASFTVLYIIOBRVNAKX	1707
Qy	1568	LQLMGGLSPITYLWGNFLWDMCNVLYPACTIVLILFLAQOQRAYVAPANTLALLLLLYG	1627
Db	1708	LQFVSGSPITYLWNTFLWDMYVNSAALVVGIFXGQKATSPENLPAITALLMLYG	1767
Qy	1628	WSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELFSQ-KLOEVSRLKQ	1686
Db	1768	WAVIPMYPASFLPDVPSTAYVALSCANLFIGINSSAITFILEFENNWTLLRFNAMLRK	1827
Qy	1687	VFLIFPHPCIGRGLIDMVRNOAMADAPERLQDQFQSPLEWVVGKILLAMVTOGFLPLL	1746
Db	1828	LLIIFPHPCIGRGLIDIALSQAVTDVYARFGEBSHTNPQWDJIGKNLVAMAAEGVVYLL	1887
Qy	1747	FTLLQHR---SOLLPOPRVRSPLLICEEDEDVARERVRVQCATQGDVLVRLNLTQVVR	1803
Db	1888	LTLFIQHHFFLTRVWSPEAKE-PII-DEDDVAEERQRIISGNKTDILRLNELTKIYS	1944
Qy	1804	QORMPADVRLCLGIPPECEFLGLGVNGAGKTSFTRMVWTDTLASRGEAVLAGHSVAREPS	1863
Db	1945	GTSSPAVDRLCVGVRPECEFLGLGVNGAGKTTFKMLTGDVTWTSGDATAGKSILTNIS	2004
Qy	1864	AAHLSMCGYCPQSDAIPELLTGREHLELLARLGRVPEAOVAQTAGSLARLGLSWYADRA	1923
Db	2005	DVHQSMGYCQCFDAVDDLLLTGREHLYIYALRGVPADEIRVANWISIQSLGLSLYADRLV	2064
Qy	1924	GTYSGGNKRKLATALALVGDPAVVFLEPTTGMDDPSARRELWNLSLAVREGSVMLTSH	1983
Db	2065	GTYSGGNKRKLSTAILMGMCPPLVLLDEPTTGMDDQARRMLWNTIYSIIREGRAVVLTS	2124
Qy	1984	SMECEALCSRLAIVWNGRCFLGSPHLLKGRFAAGHTLTLRVPAARS-----QPAAAF	2037
Db	2125	SMECEALCSRLAIVWNGRTQOCLGTIOHLKYKFGDGYIVTMKITSKPKEDLLPLDNLV	2184
Qy	2038	VAAEFPGSELRHAHGGRLRQLPPGRCALARYVGEIAVHGAESHGVDEFSVSOTMLEEVF	2097
Db	2185	FQGNFGSVQBRERNYMLQFQV---SSSSLIARIFRLIISHKDSILLEEYSVTQTLDOVF	2241
Qy	2098	LYFSKQGGKQED	2109
Db	2242	VNFAKQOTETHD	2253

This Page Blank (usp...)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:24:44 ; Search time 27.2021 Seconds
(without alignments)
5119.750 Million cell updates/sec

Title: US-09-995-542-6

Perfect score: 10896

Sequence: 1 PPLEHHECHFPKPLPSAGT.....QHPKRVQFLDDPSTAETVL 2100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters.: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5610.5	51.5	2261	4	US-09-526-193A-1
2	5603.5	51.4	2261	4	US-09-032-438C-118
3	4808	44.1	2273	4	US-09-032-438C-3
4	4730	43.4	2235	4	US-09-032-438C-6
5	3818.5	35.0	1375	3	US-08-665-259-26
6	3818.5	35.0	1375	3	US-08-762-500-26
7	2694.5	24.7	1457	3	US-08-665-259-27
8	2694.5	24.7	1457	3	US-08-762-500-27
9	2693.5	24.7	1472	4	US-09-032-438C-119
10	2507	23.0	1704	4	US-09-032-438C-120
11	2502	23.0	1704	3	US-08-762-500-75
12	2501.5	23.0	1684	3	US-08-665-259-25
13	2501.5	23.0	1684	3	US-08-762-500-25
14	427	3.9	328	4	US-09-724-797-8
15	414	3.8	607	4	US-09-252-991A-18351
16	393.5	3.6	588	4	US-09-489-039A-13579
17	389	3.6	589	4	US-09-328-352-7592
18	375	3.4	594	4	US-09-543-681A-5528
19	374.5	3.4	308	4	US-09-602-787A-348
20	371.5	3.4	335	4	US-09-252-991A-20837
21	363.5	3.3	292	4	US-09-602-787A-352
22	357.5	3.3	271	4	US-09-602-787A-350
23	354.5	3.3	315	4	US-09-328-352-4388
24	350.5	3.2	788	4	US-09-252-991A-28171
25	335	3.1	317	4	US-09-489-039A-10626
26	333.5	3.1	929	4	US-09-252-991A-22946
27	332	3.0	291	4	US-09-107-532A-4205

```

28 328 3.0 309 4 US-09-252-991A-21204 Sequence 21204, A
29 327 3.0 323 4 US-09-489-039A-12496 Sequence 12496, A
30 319 2.9 922 4 US-09-489-039A-8938 Sequence 8938, Ap
31 316.5 2.9 315 4 US-09-134-000C-6449 Sequence 6449, Ap
32 315.5 2.9 345 4 US-09-252-991A-31957 Sequence 31957, A
33 314.5 2.9 532 4 US-09-543-681A-4646 Sequence 4646, Ap
34 309 2.8 248 4 US-09-710-279-3218 Sequence 3218, Ap
35 306.5 2.8 316 4 US-09-543-681A-6184 Sequence 6184, Ap
36 301 2.8 1280 2 US-08-583-276-19 Sequence 19, Appl
37 300 2.8 1280 4 US-09-767-594-2 Sequence 2, Appli
38 300 2.8 1280 4 US-09-672-810-2 Sequence 2, Appli
39 300 2.8 1280 4 US-09-672-810-5 Sequence 5, Appli
40 300 2.8 1280 4 US-09-672-725C-7 Sequence 7, Appli
41 300 2.8 1280 6 5206352-4 Patent No. 5206352
42 300 2.8 1283 4 US-09-672-810-4 Sequence 4, Appli
43 297 2.7 391 4 US-09-252-991A-20275 Sequence 20275, A
44 297 2.7 1280 2 US-08-752-447-2 Sequence 2, Appli
45 297 2.7 1280 3 US-09-316-167-2 Sequence 2, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-526-193A-1
; Sequence 1, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pinstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526.193A
; CURRENT FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-526-193A-1

```

```

Query Match 51.5%; Score 5610.5; DB 4; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;

QY 1 PPLEHHECHFPKPLPSAGTVPWLQGLICNVNNTCFPOLTFGEFPGRLSNFNDLSVRL 60
Db 47 PPYEHECHFPKAMPAGTLPWVQGIICNANPCFRYPTEGAPGVGVGNFKSIVARLF 106
QY 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAQ-----PQ 96
Db 107 SDARRLLYSOKDTSMKMRKVLRTLQIKKSSNKLQDFLVNDETSGFLYHNLSLPK 166
QY 97 PTKQSPLEPPML-----DVAEL----- 113
Db 167 STVDKMLRADVILHKVFLQGVQLHLTSLCNGSKSEMIQLGDQFVSELGFLPREKLAAR 226
QY 114 -----LTSLLRT--ESIGLALGQAQEPHLSLEAAEDLAQELALRSLEVL--A 159
Db 227 RVLRSNMDILKPIILRTLNSTSPFPFSKELAEATKTLHSLGTLAQELFSMRSDMRQVM 286
QY 160 LLQPRGTSGLPEL---LSEALCSVRGSPSTVGFSLNWEASDLMELVG----QEPESAL 212

```


Db 1756 ASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTNNKLNLDINDILKSVFLIPFHECL 1815
Qy 1651 GRGLIDWVRNQAMADAFERIGDRQFQSPRLRWEVVGKLLAWIOGPIFLFTLLQHRSQ 1710
Db 1816 GRGLIDWVRNQAMADALERFGENFVSPSLNDLVRNLFAVAVGVVFFLITVLIOYRFF 1875
Qy 1711 LLPQPRVRSPLLLGEEDEVDARERVRVQAGATQGDVLVRLNLTQVYRGQMPADVRLCLG 1770
Db 1876 IRPRPVAKLPPLNDEDEDVRRERQRIIDGGGQNDILEIKELTKIYRRKRKPAVDRI:G 1935
Qy 1771 IPGCEGFLGVNAGAGTSFPMVTGDTLASRGEAVLAGHSVAREPSAAHLGSMGYCQPSD 1830
Db 1936 IPGCEGFLGVNAGAGTSFPMVTGDTLASRGEAVLAGHSVAREPSAAHLGSMGYCQPSD 1995
Qy 1831 AIFELLTGREHLELLARLGRVPEAQVATAGSGMARLGLSWADRPAGTYSGGNKRKLAT 1890
Db 1996 AITELLTGREHVEFFALLRGVPEKEVKGFGWAIKGLGVKGYKASNYSGGNKRKLST 2055
Qy 1891 ALALVGDPVAVFLDEPTTGMDPSARRFLWNSLLAVVREGSRVMLTSHSMBECEALCSRLA 1950
Db 2056 AMALIGPPVAVFLDEPTTGMDPKARFLWNCALSIVKGRSVVLTSHSMBECEALCTRMA 2115
Qy 1951 IMVNGRFRCLGSPHOLKGRFAAGHTLRLRPAARS--OPAAAFAVAAEFPGSELREAHGGR 2008
Db 2116 IMVNGRFRCLGSPHOLKGRFGDGTIIIVRIAGSNPDLPKVOEFPGELAPFGSVLKEKRN 2175
Qy 2009 LRFOLPFGGRCLGARVFGELAVHGAEGVEDFSVQTMLEEVFLYFSQDQKDE--DTE 2065
Db 2176 LOYQL-FSSLSLARIIFSILSQSKRLHIEDYSVSTQTLDOVFVFAKQSDDDHLKDL 2234
Qy 2066 EKEAGGVDPAPGLQHPKRVSVQFLDDPSTAE 2098
Db 2235 LHKNTQTV-VDVAV-----LTSFLODEKVKES 2259

RESULT 3
US-09-032-438C-3
; Sequence 3, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; FILE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 2273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-3

Query Match 44.1%; Score 4808; DB 4; Length 2273;
Best Local Similarity 44.2%; Pred. No. 0;
Matches 1007; Conservative 354; Mismatches 635; Indels 282; Gaps 39;

Qy 1 PPLEHHECHPNKPLPSAGTVPMLOGLICNVNNTCPQLTPGEBGELSNFNDSLVSRLL 60
Db 47 PLYSHHECHPNKAMPAGMLPWLQGLFCNVNNTCPQSPPTGSPGIVSNYNNISILARVY 106
Qy 61 ADARTVILGGASAHRTLAG-----LCKLIATLRA 88
Db 107 RDFQELLMNAPESQHLGRITWELHILSOFMDTLRTHPERIAGRIRDIRIDILKDEETLTLF 166
Qy 89 -----ARSTAQPOPTKQSPLEPPMLDVA----- 111
Db 167 LIKNIGLSDSVVYLLINSQVRPEQFAHGVDPDLAKDIACSEALLERFIIFSQRRAKTVR 226
Qy 112 -----ELTSLRTESTGLGALQAQEPHLSLEAAED 143
Db 227 YALCSLSQGTLOWIEDTLVANVDFPKLFRVLPILTLLDSRSQGINL-----RSWGGILSDMSP 282
Qy 144 LAQELLALRSLVEL-----RALLQR--PRGTSGLPELLSEALCSV-RGFSSTVGPSPNLNWE 196
Db 283 RIQEFTHRPSMQDILLWTRFLMONGPETFTKLMGILSDLLCGYPGGGSRV-LSFNWYE 341
Qy 197 ASDLMELVGOEPESALP-----DSSLSPACSELIGALDSHPLSRLLMRRRLKPLILGLKLLFA 252
Db 342 DNNVKAFLGIDSTRKDIYSYDRRTTSFCNALIQSLESNELTKIAWRAAKPFLMGKILYT 401
Qy 253 PDTPTFRKLMAQVNRTEFEELTLRDVREWEMLGPRIFTFMNDSSNVAMLQRL----- 306
Db 402 PDSAPARRILKANANSTFELEHVRLKVAWEVGPOIWFYFDNSTQNMNRDITLGNPTVK 461
Qy 307 -----QWODEGRQRPGRGRDHMEALRSFLDPG-----SGGYSMDAHADVHVLGT 353
Db 462 DFLNRQLGEGIT-----AEALNFIYKGPRESQADDMANFDWRDIFNITDRTLRL 512
Qy 354 LGRVTECLSLDKLEAAPSEAAALVSRAQLLAERHFWAGVVLGPESSDTEPTDPLGP 413
Db 513 VNQYLECLVLDKPESYNDETQLTQRALSLLEENFWAGVVF-----DMYPTWSSLP 564
Qy 414 GHVRIKIRMDIDVYTRNKIRDRFWDGPAADPLTLRLYVYVGGFVYLODVERAAVRVLS 473
Db 565 PHVKYKIRMDIDVVEKTKIKDRYWSGPRADVEDFRYIWGGFAYLQDWMVEQGITRSQV 624
Qy 474 GANPRAGLYQOMPYPYCVDDVFLRSLRSLPLFLTLAWIYSVTLTKAVVREKETLRD 533
Db 625 QAEAPVGIYLOMPYPCFVDDSDFMILNRCFFIFMWLAWIYSVMTVKSIVLEKELRKE 684
Qy 534 TMRAMGLSRVAVLWGLWFLSCGLPPLLSAALLVLVLKGLDILPYSHPGVVFLLFAFVAT 593
Db 685 TLKNQGVSNVAVIWTWFLDSFSIMSMISIFLLTIFIMHGRILHYSDPFIPLFLILAFSTAT 744
Qy 594 VTQSFLLSAFFSRANLAAACGGLAYSLYLPYVLCVAVWRDLRDPAGRVAASLSPVAFGF 653
Db 745 IMLCFLLSLSTFFSKASLAAACSGVIYFTLYPLHLCFAWQDRMTAELKAVSLSPVAFGF 804
Qy 654 GCSLALLBEOGGAQWNNVGTPT-ADVFSLAQVSGLLILLDLYLATWYLAIVCAVPGQ 712
Db 805 GTEYLVRFEGQGLGLONSIGNSFTGDEFSFLLSQMMLDAAVYGLLAWYLDQVFGD 864
Qy 713 YGIPPMNPFRRSYMGCRPPPKSPAPCPT-----PL-----DPK-----VLV 750
Db 865 YGTPLPWYFLQBSYVLGGE-----GCSTREERALEKTEPLETEEDPEHGHDSFF 918
Qy 751 BEAPPGLSPOVSYSRSLSEKRPFGSPQPALRGLSLDFYQGHITAFIGHNAGKTTTLTSLG 810
Db 919 EREHPGVVPGVCVKNLVKIIEPCGRPAVDRLNITFYENQITAFIGHNAGKTTTLTSLT 978
Qy 811 LFPSPGGSFALIGHDVRSSMAAIRPHLGVCPQVNVLFMDLTVDEHVMFYGRKLKSAV 870
Db 979 LLPPTSGTVLVGRDIETSLDVAQRSLGCMCPQHNLPHHLTVAEHMLFYAQLKGSQBEA 1038
Qy 871 GPEQDRLLQDVLVSQSVQTRHLSSGMQRKLSVAIAFVCGSQVILDEPTAGVDPASRR 930
Db 1039 QLEMEAMLEDTGLHHKRNESAQDLSCGMQRKLSVAIAFVGDVAKVILDEPTSGVDPSRR 1098
Qy 931 GIWELLKYREGRTLILSTTHLDEALLGLDRVAVVAGRLCCCGSPFLFRHLLSGSYLT 990

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1375 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-665-259-26

Query Match 35.0%; Score 3818.5; DB 3; Length 1375;
Best Local Similarity 54.1%; Pred. No. 0;
Matches 754; Conservative 224; Mismatches 348; Indels 69; Gaps 14;

QY 750 VEEAPPGLSPGVSRSLKRPSPQPALRGLSLDFYQGHITAFIGHNGAGKTTLSILS 809
DB 2 MEEETHRLRLGVSIGNLVKVRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSILT 61
QY 810 GLFPSPGSAFLGHHDVRSMAAIPHLGVCPQYNVLFDMLTVDHVFYGRKLSAAV 869
DB 62 GLFPPTSGTAYTLGHDIRESMSIRONLGVCPQHNVLFDMLTVEHIFWYARLKLSEKH 121
QY 870 VQPEODRLLODVL--VSKQSVOTRHLSGMOKLSVAIAFGVGSQVWILDEPTAGVDPAS 928
DB 122 VKAEMQALDVGLPSPKSKTSQSGMOKLSVALAFVGGSKVIVLDEPTAGVDPIYS 181
QY 929 RRGWELLKYREGRTLILSTHLLDEALLGDRAVAVAGGRLCCCGSPFLRRLHLSGYY 988
DB 182 RRGWELLKYRQRTIILSTHMDADILGDRIAISHGKLCVGGSSFLKQNLQGTGY 241
QY 989 LTLVKARLPLTN-----EKADTMEGSDVTRQEKNGSQSRVGTPTOLLALV 1036
DB 242 LTLVKDYESSLSCRNSSSTVSCLLKEDSVSQSSDAGLGDHESDTLTIDVSAISNLI 301
QY 1037 QHWVPGARLVEPLHELVLPLPYTCAHDSFATLFLRELDTLRLAEURLTGYGSDTSLEI 1096
DB 302 RKHVSEARLVEDIGHELTVLPYEAKEGAFVELFHEIDRLSDLCISYGISSETLBEI 361
QY 1097 FLKVVCECAADTMDG-----SCQHLCTGTAGDVTLRLLKMPQETAL 1141
DB 362 FLKVAEESGVDAETSDGTLPARNRRAFGDKQSC-LHPFTEDDAVD-----BNDSDI 412
QY 1142 ENGEPAGAPETDQSGDPAVG--RVQGWALTRQOLQALLKRLARRSRGLFAQIVL 1199
DB 413 D-----PESRETDLSSGMDGKGSYQLKGWKLTKQOQFVALLKRLIARRSRKGFPAQIVL 467
QY 1200 PALFVGLALVSLIIVPPGHVPAIRLSPTMYGAQVSFFSEADAPGDPGRARLLEALLQBAQ 1259
DB 468 PAVFVCIALLVSLIIVPPGKTPSLLEQPMWNEQTTFVSNDAPEMDGTQELLNALTQDPG 527
QY 1260 -----LBEPPVQSHSRHFAPEVAEYAKVLASGNMTPEPSFACQCSQFGARRL 1309
DB 528 FQTRCMEGNPIPDFTCLAGEEDWTISVPQSVLDLFGNGWMTKPNPSFACQCSQSKIKKM 587
QY 1310 LPDCAAGAGPPPPQAVTGSBQVNLGRLNLSDFLVKTYTFLVROGLTKKWNVEVRYG 1369
DB 588 LVPVCPGAGGLPPPPQKOKTADILQNLGRLNISDYLVKTYTQIIAKLSLKNKIWNVEFRYG 647

QY 1370 QFSLG-GRDPGLPSQELGRSVBELWALLSLPLPGGALDRVLKNTLTAHSLDAOBSLKW 1428
DB 648 QFSLGVNSQALPPSHEVNDAIKQMKLLKLTQDTSADRFLSSLRFPAGLDTKNVKNV 707
QY 1429 FNNKGWHSMAFVNRASNAIIRAHLPQPARHAHSITTLNHPNLTKQSLSEAAALWASSV 1488
DB 708 FNNKGWHAISSFLAVINNAIRANLQKGENPSQYGITAFNHPNLTKQQLSEVALMTTSV 767
QY 1489 DVLVSICVVFVAMSFPASFTLVLTBEERVTRAKHLQMLGGLSPTLVYMLGNFLWDMCNLYVP 1548
DB 768 DVLVSICVVFVAMSFPASFTLVLTBEERVTRAKHLQMLGGLSPTLVYMLGNFLWDMCNLYVP 827
QY 1549 ACIVVLIFLAFQORAYAPANLPAIILLLLLLYGWSITPLMYPASFFSVSPSTAYVLTIC 1608
DB 828 ATLVIIFICFQKQSVGSSTNLVLAALLLYGWSITPLMYPASFFSVSPSTAYVLTISV 887
QY 1609 NLFINGSMATFVLELPSDOKLOEVSRIILKQVFLIFPHFCLGRGLIDMVRNQAADAFE 1668
DB 888 NLFINGSVATFVLELFTNNKLANDINDILSKSVLIFPHFCLGRGLIDMVRNQAADAFE 947
QY 1669 RLGDROFOSPLRWVGNKLLAMVIOGFLFLLTLLLOHRSOLLPOPRVRSPLLGERDE 1728
DB 948 RFGENRFVSPSLWDLVGRNLFAMAVEGVVFFLITVLIQYRFFRPRVYAKLPLPNDSE 1007
QY 1729 DVARERERVQATQGDVLVLRNLTKVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKT 1788
DB 1008 DVRRERQRILOGGQNDILEIKELTKIYRRKRKPAVDRAICIGIPPGECFGLLVNGAGKS 1067
QY 1789 STFRMVTGDTLASGEAVLAGHSVAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARL 1848
DB 1068 TTFKMLTGDTPVTRGDAFLNKNLSILSNIEHVQNMGYCQPDATITELLTGREHVEFFALL 1127
QY 1849 RGVPBAOVAQTAGSLARLGLSWADRPAQTYSGNKEKLTALVGDPAVFLDEPT 1908
DB 1128 RGVPKEVKGEGEWAIRKGLGVKYGKYSYSGNKRKLTAMALIGGPPVFLDEPT 1187
QY 1909 GMDPSARFLWNSLLAVVREGSVMLTSHSMECEALCSRLAIVNGFRCLGSPQLKLG 1968
DB 1188 GMDPKARRFLWNCALSIKVEGRSVVLTSHSMECEALCTRMALVWNGFRCLGSPQLKLG 1247
QY 1969 RFAAGHTLTLRVPAARS--QPAAPFAABFPGSELREAHGRLRPLQPPGRCALARVFG 2026
DB 1248 RFGDGYTIVRIAGSNPDLKPVQSFPGFGLAFPGSVLKEKRNMLQYL--PSSLSLARI 1306
QY 2027 ELAVHGAHGVDESVSOTMLEEVFLYPSKQDQDE---DTEEQKEAGVDPAPGLQHP 2083
DB 1307 ILSQSKRLHIEDYSVSQTTLDQVFNPAKQSDDDHLKDLHLKNTQTV--VDVAV----- 1360
QY 2084 KRVSQFLDDPSTAET 2098
DB 1361 --LTSFLQDEKVKES 1373

RESULT 6
US-08-762-500-26
Sequence 26, Application US/08762500
Patent No. 6030806
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America

NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1457 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-665-259-27

Query Match 24.7%; Score 2694.5; DB 3; Length 1457;

Best Local Similarity 40.8%; Pred. No. 4.8e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

QY 750 VEEAPPGLSPGVSRSLKRPSPQALRGLSLDFYQCHITAFIIGHNGAGKTTLSLS 809
DB 1 MEEETHPLVVCVDKLTQYKNDKLLNLNLYENQVVSFLHNGAGKTTMSILT 60
QY 810 GLFPSPGSAFLGHDVSRSSMAIRPHLGVCQYNVDFMLTVDHVMFYGRKLGLSAAV 869
DB 61 GLFPPTSGSATYGHDIRTEMDEIRKNLGMCPOHNVLFDRLTVEHLPYSLKSAQEE 120
QY 870 VCPEQDRLLQDVLGYSKQSVQTRHLSGGMOKLSVAIAFVGSQVVIDEPTAGVDPASR 929
DB 121 IRKETDKMIEDLSNKRHSVLQTLSSGGWKRKLSVAIAFVGSRAIILDEPTAGVDPYAR 180
QY 930 RGIWELLKRYREGRTLILSTHLDDEALLGDRVAVVAGRLCCGSPPLFRHLGSGYYL 989
DB 181 RAIWDLIILKYKGRITLSTHMDADLLGRIALISHGLKCCGSPFLKAGYXDGRL 240
QY 990 TLVKARLPLTNEKADTMEGSDVTRQEKNGSQSGSRVGTQQLALVQHVPGARLVEEL 1049
DB 241 TLVKQPAEPGTQSEGLASSPGCPRL-----SSCSEPVQSQFIRKHVASSLLVSDT 292
QY 1050 PHELVLVLPYGAHNGSATPLFRELDTLAEURLTGYGIDSTLSBEEIFLKVEE----- 1103
DB 293 STELSYILPSEAVKKGAFERLQFQLEHSLDALHLSFGMLDPTLBEVFLKYSSEEDQSLEN 352
QY 1104 CAADT-----DWDGSCGQHLCTGAGLDVTLRLKAMPQETALENGEPAGSAPETDQSGG 1158
DB 353 SEADVKEKRDVLPAGELTAVGGQAG-NLARCSELAQSASLSQASSVGS-A-REEGTG 410
QY 1159 -----PDVAG-----RVQGWALTQQQLALLKRPL 1184
DB 411 YSDGYDYRPLFDNLQDPDNVSLQAEAMEALAQVQSGSKLEGWMLKVRQFHLVKKRFH 470
QY 1185 LARRSRGLFAQIVLPALFVGLALVFSLLVPPFGHYPALRLSPTWY-----GAQVSFF 1237
DB 471 CARRNSKALCSQILLPAPFVVCVAMTVALSVPEIGDLPLVLSPSQYHNYTPQPRGNFIPIYA 530

RESULT 8

US-08-762-500-27

; Sequence 27, Application US/08762500

; Patent No. 6030806

QY 1238 SED-----APGDPGRARLLEALLQAG-----LBEP-----PVQSSH----- 1270
DB 531 NEEROEYRLRLSP-DASQQLVSTFRLPSGVGATCVLKSPPANGSLGPMNLNLSGSRLLA 589
QY 1271 -RFSAPVPAEYAKVLASGNWTPESPAA----- 1298
DB 590 ARFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVKXPDEDSLQAWNMSLPPTAGPETWTS 649
QY 1299 -----COCQOPGARRLLPDCPAAAGGPPPPQAVTGSSEVQVNLTRNLSDPL 1345
DB 650 APSLRLVHEPVRCTCSAQTGF-----SCFSSVGGHPPQWRVV-TGDILTITGHNVSEYL 705
QY 1346 VKTVPRLVRQGLKTKKWNVRYGFSLGRDPLPS--GQBLGRSVIELWALLSPLPGG 1403
DB 706 LFTSDRFLH-----RYGATITGNVQKSIIPASFGARVPPWVKI----- 744
QY 1404 ALDRVLKNTAWAHSLSDAQDSLKIWFNNKGMHSMVAVPNRASNAIILRAHLP--GPARRH 1460
DB 745 AVRRV-----AQVLYNNKYHSMPTYLNSLNLNLAIRANLPSKGNPA-- 786
QY 1461 AHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVFPAMSFVPASFTLVLIIEERVTRAK 1520
DB 787 AXITVTNHPMNTKSASLSLYLL-QGTDVVIAPFIIVAMSEVPASFVVFVLAESTKAK 845
QY 1521 HLQLMGSLPTLYMLGNFLWDMCNVLPACIVLIFLAFQOQRAYVAPANLPAALLLLLY 1580
DB 846 HLQVSGCNPIVYMLNVWDMNLVLPATCCVILFVFDLPAYTSPNFPVAVLSFLY 905
QY 1581 GWSITPLMPASFPFSPSTAYVVLTCINLFIGINGSMATPVLELFS-DQKLQVSRILK 1639
DB 906 GWSITPIWYASFWPEVPSAYVFLVNLFIGITATVATFLLQFELDKDLKVNSYLK 965
QY 1640 QVELIFPHFCGLGRGLDMVRQAMADAFERLGD-RQFOSPLRWEVVGKXLLAMVLCQPLF 1698
DB 966 SCFLIFPNYINLGHGLEMAYNEYIYAKIQGQDKMSPFEDWITVTRGLVAMTVGEFVG 1025
QY 1699 LLFTLLQLHRQLLPQR---VRSPLIGEDEDVARERERVVQAGTQGDVLVLRNLTKV 1755
DB 1026 FFLTIMQY--NFLRQFORLPVSTKPV--EDDVVASERQVRLRGDADNDVWKIENLTKV 1081
QY 1756 YRGO---RMPAVDRLCIGI-PPGECFGLLVNGAGKTTFRMVGTDTLASRGEAVLAGHS 1811
DB 1082 YKSRKIGRILAVDRLCIGVCVPGECFGLLVNGAGKTTFRKMLTDESTTGGEAFVNGHS 1141
QY 1812 VARPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLABGLSM 1871
DB 1142 VLKDLQVQQLGVCPCPDVDPVDELTAHEHLQLYTLRCIPWKDBAQVVKWALEKLETK 1201
QY 1872 YADRPAGTYSGNKRKLATALLALVGDPAVAVFLDEPTTGMDFPSARRFLWNSLIADVREGS 1931
DB 1202 YADRPAGTYSGNKRKLSTALALGYPAFIFLDEPTTGMDFPSARRFLWNLILDLIKTGRS 1261
QY 1932 VMLTSHMEECEALCSRLAIMVNGRFRCLGSPQLHKGFAAGHTLTLRVPAARS-OPAAA 1990
DB 1262 VMLTSHMEECEALTRLAIMVNGRLHCLGSIQHLKRNFGDGYMITVTKSSQNVKDVVR 1321
QY 1991 FVAEFPCESELRBAHG-----RLRFPQPGRCALARVFGELAVHGAHGVDEFSVQ 2045
DB 1322 FPNRNP-----EHAQOKTYPVQYQL-KSEHLSAQVFSKMEQVGVGLGIEDYSVQ 1375
QY 2046 MLEEVLYFSKQKQKDETEQKAGVGVDPAP-----GLQHPKRVSOFL-----DDPSTA 2096
DB 1376 TLDNVFVNFAK---KQSDNVEQOEAPSSLSPLGLLLSLPRPAPTELRAVVADEPEDL 1432
QY 2097 ET 2098
DB 1433 DT 1434

GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1457 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-762-500-27

Query Match 24.7%; Score 2694.5; DB 3; Length 1457;
Best Local Similarity 40.8%; Pred. No. 4.8e-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;

QY 750 VEEAPPGLSPGVRSLEKRPSPQALRGLSLDFYQGHITAPLGHNGAGKTTTLTILS 809
DB 1 MEEBETHPLVVCVDKLTQVYKNDKLANLKLNLNLYENQVVSFLGHNGAGKTTMTSILT 60

QY 810 GLFPSPGSAPILGHVRSMAAIPHLGVCPOQNVLFMDLTVDEHWFVGRKLGLSAV 869
DB 61 GLFPPTSGSATYGHDIRTEMDEIRKNGLMCPQHNVLFDRLTVBEHLWFSRLKSMQEE 120

QY 870 VGPEODRLQDVGIVSKQSVQTRHLSGMORKLSVAIAFVGSQVWILDEPTAGVDPASR 929
DB 121 IRKETDKMIELESNKRHSIVQLTSGMKRKLVAIAFVGSRAIILDEPTAGVDPYAR 180

QY 930 RGIWELLKLYREGRTLILSTHLLDEAELLGDRVAVAGGCCGSPFLFRRHILGSGYL 989
DB 181 RAIWDLILKYKPGRTILSTHMDADLGDRIAIISHGLKCCGSPFLKGYXDGRL 240

QY 990 TLVKARPLLTNEKADTMDEGSDVTRQKNGSGGSRVGTGPOLLALVOHWVPGARLVEEL 1049
DB 241 TLVRQAPBPGTSQBPGLASSPGCPRL-----SSCSEPQVQSFIRKHWASSLLVSDT 292

QY 1050 PHELVVLVLPYTGADGGSPATLRFRELDTRLAELRLTGVGISDTSLEEIFLKVVEE----- 1103

Db 293 STLSYILPSEAVKKGAFERLFOOLEHSLDALHSSFLGMDTTLLEEVFLKVEEDQSLEN 352
QY 1104 CAADT-----DMEDGSCGHLCTGIAGLDVTLRLKMPPEQTALENGEPAGSAPETDQSG 1158
Db 353 SEADVKSERKDVLPGAEGLTAVGQAG-NLARCELAQSQASQSSVSSA-RGEGTG 410
QY 1159 -----PDVAG-----RVQGWALTRQQLQALLKRFL 1184
Db 411 YSDGYGDRPLFDNLQPDNVSLSQAEAMEALAQVQGSRLKLEGWLMKMRQPHGLLVKRFH 470
QY 1185 LARRSRGLFAQIVLPALFVGLALVFSLIYPPGHYPALRLSPMY-----CAQVSFF 1237
Db 471 CARRNSKALCSQILLPAFFVCVAMTVALSPEIGDLPLVLSPSQYHNYTQPRGNFIPYA 530
QY 1238 SED-----APGDPGRARLEALLQEAG-----LEEP-----PVOHSSH----- 1270
Db 531 NEERQEVRLRLSP-DASPOQLVSTFRPLPSGVGATCVLKSPPANGSLGPMNLSSGESRLA 589
QY 1271 -RFSAPVPAEVAKVLASGNWTPESPSPA----- 1298
Db 590 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVXPDEDSLOANWMSLPPTAGPETWTS 649
QY 1299 -----COCSPGARLLPDCPAAAGPPPPQAVTGSVEVQNLTRNLSDFL 1345
Db 650 APSLPLRVHEPVRTCQAQGTG-----SCPSVSGHPQMRVV-TGDLTITGHNVSEYL 705
QY 1346 VKTYPRLVROGLTKKWNVEVRYGGFSLGGRDGLPS--QOELGRSVSEELWALLSPLPGG 1403
Db 706 LFTSDRFLH-----RYGAIITGNVQKSIIPASFGARVPPMVKI----- 744
QY 1404 ALDRVKNLTAWAHSLLDAQSLKIFWNNKWHSHVAPVNRASNAIILRAHLP-----GPAPH 1460
Db 745 AVRRV-----AOVLNNKGYHSMPTYLNSLNAILRANLIPKSKGNPA-- 786
QY 1461 AHSITTLNHLPLNLTKQOLSEAAALMASSVDVLVSCVVFAMSFVPASFTLVLIERVTRAK 1520
Db 787 AYXITVTNHPMNTKSLSLDYLL-QGTDVVIATFIIVAMSFVPASFPVFLVAKSTKAK 845
QY 1521 HLQLMGGLSPTLYWGLNFMCMNYLPACIVVLIFLAFQORAYVAPANLPALELLLLY 1580
Db 846 HLQFVSGCNPIVYLANVYVMDLNYLPATCCVILPVFPLPAYTSPNTPAVLSFLLY 905
QY 1581 GWSITPLMYPASFPSPSTAYVVLTCINLFIGINGSMAFVLELFS-DQKLQVSRILK 1639
Db 906 GWSITPTMPASFWEFPSSAYVFLIVINLFIGITATVATVATLQLFHDKDLKVNSYLK 965
QY 1640 QVFLIPPHFCLGRGLIDMVRNOAMADAFERLGD-RQFQSPLRWEVVGKLLAMVIOGPLF 1698
Db 966 SCFLIFPNYVNLGHLMEAMAYNEIYAKIQDPKMKSPFWDIVTRGLVAMTVGEFVG 1025
QY 1699 LLFTLLQHRSQLLPQR---VRSPLLLGDEEDVARERERVQOGATGQDVLVLRNLTKV 1755
Db 1026 FFLTIMQY--NFLRQQRPLPVSTKPV--EDDVVASERQVRLRGDADNDNMVKIENLTKV 1081
QY 1756 YRGO---RMPAVDRCLGI--PPGECFGLLVNGAGKTTSTFRMTGDTLASRGEAVLAGHS 1811
Db 1082 YKSRKIGRIILAVRLCLGVCPGECFGLLVNGAGKTTSTFKMLTGDESTTGGEAFVNGHS 1141
QY 1812 VAREPSAHLISMGYCPOSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSW 1871
Db 1142 VLKDLQVQSLGYCPOFDVPVDELTAHQLYTRLCRCPKDEAQVVKWALEKLELTK 1201
QY 1872 YADRPAGTYSGNKRKLATATALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVRGRS 1931
Db 1202 YADKPAGTYSGNKRKLSTAIALIGYPAFLDEPTTGMDPKARRFLWNLILDIKTGRS 1261
QY 1932 VMLTSHSMECEALCSPLATMNGRPFCLGSPHLLKGRFAAGHTLTLRVPAARS-QPAAA 1990
Db 1262 VMLTSHSMECEALCTRLAIVNGRLHCLGSIQHLKRRFGDGYMITVTKSSQNVKQDVR 1321
QY 1991 FVAAEFFGSELREAHGG-----RLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVSQT 2045

Db 1322 FFRNFP-----EAAHQKTPYKVQYL-KSEHISLAQVFSKMEQVGVGLGIEDYSVSQT 1375
Qy 2046 MLEEVFLYFSKQDQEDTEQKEAGVGDPA-----GLQHPKRVSQEL-----DDPSTA 2096
Db 1376 TLDNVFVNFAK---KQSDNVEQOEAPSPSLPSLGLLSLLRPRPAPTRELRAVADEPDL 1432
Qy 2097 ET 2098
Db 1433 DT 1434

RESULT 9

US-09-032-438C-119
; Sequence 119, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupeki, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 1472
; TYPE: PRT
; ORGANISM: Mouse
; US-09-032-438C-119

Query Match 24.7%; Score 2693.5; DB 4; Length 1472;
Best Local Similarity 40.8%; Pred. No. 66-231;
Matches 613; Conservative 244; Mismatches 424; Indels 221; Gaps 37;
Qy 750 VEEAPPGLSPGVSVRSLEKRPFGSPQPALRGLSLDFYQGHITAFILGHNGAGKTTTLSTLS 809
Db 16 MEEEPHTLPLVVCVDKLTQVYKNDKLAALKSLNLYENQVSVFLGHNGAGKTTTMSILT 75
Qy 810 GLFPFGSGSATILGHVRSMAARPHGVCPQYNNVDFMLTVDHVVYFGRKGLSAAV 869
Db 76 GLFPPTSGSATYGHDIRTEMDEIRKNGMCPQHNVLFDRLTVEHLWFYSLKSMQEE 135
Qy 870 VGPEODRLLODVLVSKQSVOTRHLSCGMQKLSVAIAFVGSQVWILDEPTAGVDPASR 929
Db 136 IRKETDKMIEDELSNKRHSILVQTLSGGMKRLSVAIAFVGSRAIILDEPTAGVDPYAR 195
Qy 930 RGIWELLIKYREGRTLILSTHLLDEALLDRVAVVAGRCCLCGSPFLFRLHLSGYYL 989
Db 196 RAIWDLILKYKGRITLSTHMEADLLGRIALISHGKLLKCCGSPFLFGAYKDGURL 255
Qy 990 TLVKARLPLTTNEKADTMEGSDVTRQEKNGSQGSRVGTPTQLLALVQHWVPGARLVEEL 1049
Db 256 TLVQAPBPGTSQBPGGLASSPGCPRL-----SSCSEFQVSQFIRKXVASSILLVSDT 307
Qy 1050 PHELVLVLPYTGANDGSPATIFRELDTRFASRLRTGYSIDTSLEETPLKVVE----- 1103
Db 308 STELSYILPSAVKKGAFERLFOLEHSLDALHLSFGLMDTTLVEEFLKVSSEEDQSLEN 367

Qy 1104 CAADT-----DMEDSCGCHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSG 1158
Db 368 SEADVKEERKDVLPAGEGLTAVGQAG-NLARCSELAQSOASQASSVGSAR-REBGTG 425
Qy 1159 -----PDAVG-----RVQGWALTQQQLQALLKRLP 1184
Db 426 YSDGYDYRPLFDNLQDPDNYSLQEAEMEALAQVGGSRKLEGMWLMKRPFHGLLVKRFH 485
Qy 1185 LARSRRLGFAQIVLPALFVGLALVPSLIVPPFGHYPALRLSPMTWY-----GAQVSFF 1237
Db 486 CARRNSKALCSQIILLPAPFFCVAMTVALSVEIGDLPPLVLSPSQYHNYTQPRGNFPIYA 545
Qy 1238 SED-----APGDPGRARILLEALQAG-----LEEP-----PVQHSH- 1270
Db 546 NEERQEVRLRLSP-DASQQQLVSTFRLPSGVGATCVLAKSPANGSLGPMNLSSGESRLLA 604
Qy 1271 -RFSAPVPAEVAKVLASGNWTPSPSPA----- 1298
Db 605 ARFFDSMCLESFTQGLPLSNFVPPPPSPAPSDSPVKPDEDSLQAWNMSLPTTAGPETWTS 664
Qy 1299 -----CQCSQPGARRLLPDCPAAAGPPPPQAVTGSCEVVQNLTRNLSDPL 1345
Db 665 APSLPRLVHEPVRTCSAQGTG---SCPSSVGGHPQMRVV-TGDIITDITGHNVSYL 720
Qy 1346 VKTYPRLVROGLTKKWNVEVRYGFGSLGRDPLPS--GOELGRSVEELWALLSPLPG 1403
Db 721 LFTSDRPLH-----RYGAIITFGNVKQSPASFGARVPMVRKI----- 759
Qy 1404 ALDRVLKMLTAWHSLDAQDSLKTWFNNKWHSMVAFVNRASNAILRAHLP-----GPARH 1460
Db 760 AVRRV-----AQVLYNNKYGHSMPTYLNSLNNAILRANLPSKGNP-- 801
Qy 1461 AHSITTLNHLPLNTKEOLSEALMASSVDVLVSCVVPAMSPVSPASFTVLIEERVTRAK 1520
Db 802 AYKITVTNHPMNTSASLSLDYLL-QGTDVVIAIFIIVAMSFVSPASFPVFLVAEKSTAK 860
Qy 1521 HLQLMGGLSPTLYWLGNFMDMCMYLPACTVWILFLAFQORAYVAPANLPAALLLLLY 1580
Db 861 HLQFVSGCNPIYWLAVNMDMLNLYLPATCCVILFVFDLPAYTSPTNFPAVLSFLLY 920
Qy 1581 GWSITPLMYPASFPSPSTAYVVLTCINLFIGINGSMAFVLELFS-DQKLQEVSRILK 1639
Db 921 GWSITPIMPASFPWFEPSSAYVFLIVNLFIGITATVATFLQLFEHDKDLKVNSYLK 980
Qy 1640 QVLEIFHFCLGRGLIDMVRNQAMADAPERLGD-RQFOSPLRWEVGNLAWMTQGPLF 1698
Db 981 SCFLIFPNYLNHGLMAYNEYINEYAKIQGDFMKSPFEDWITVRLGLVAMTVEGVFG 1040
Qy 1699 LLFTLLQHRSQLLPQPR--VRSPLLGEDEDAVERERERVVOGATQGDVLRNLTKV 1755
Db 1041 FFLTIMQY--NFLRQORLPVSTKPV--EDDVASERQORVLRGDADNDVMIENLTKV 1096
Qy 1756 YRGO---RMPAVDRCLGI--PPGECFGLLVNGAGKSTTFMWVGTDTLASGEAVLAGHS 1811
Db 1097 YKSKIRIILAVDRCLGVCVPGCFGLLVNGAGKSTTFKMLTGDESTTGGEAPVNGHS 1156
Qy 1812 VAREPSAAHSMGYCCOSDAIFELLTGREHLELLARLGRVPEAQVQVATAGSGLARLGLSW 1871
Db 1157 VLKOLLVQVQSLGVCPOFDPVDELTAHQLYLRICIPWKDEAQVVKWALEKLEUTK 1216
Qy 1872 YADRPAGTYSGNKRKLATALLVGDPAVPLDPTGMDPSARRFLWNSLLAVVREGS 1931
Db 1217 YADRPAGTYSGNKRKLSTALIGYPAFPLDDEPTTGMDPKARRFLWNLLDLIKTGRS 1276
Qy 1932 VMLTSHNMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARS-QPAAA 1990
Db 1277 VMLTSHNMECEALCTRLAIMVNGRLCLGSIQHLKNRFGDGYMITVTKSSQNVKDVVR 1336
Qy 1991 FVAEEFPGSELREAHGG-----RLRFOLPGRCALARVCELAVHGAHEGVDFSVSQT 2045
Db 1337 FFRNFP-----EAAHQKTPYKVQYL-KSEHISLAQVFSKMEQVGVGLGIEDYSVSQT 1390
Qy 2046 MLEEVFLYFSKQDQEDTEQKEAGVGDPA-----GLQHPKRVSQEL-----DDPSTA 2096

Db 1391 TLONVFNFAK---KQSDNVEQOABPSLPSPLGLLSLLRPRPAPTELRAALVADEPEDL 1447
Qy 2097 ET 2098
Db 1448 DT 1449

RESULT 10
US-09-032-438C-120
; Sequence 120, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h P.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 120
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-032-438C-120

Query Match 23.0%; Score 2507; DB 4; Length 1704;
Best Local Similarity 33.7%; Pred. No. 3.9e-214;
Matches 639; Conservative 314; Mismatches 630; Indels 314; Gaps 47;

Qy 261 LMAQVNTFBELTLRDVREVMELGPRIFTFMDSNVAMQLRLQOMDEGRQRPPGG 320
Db 9 LLLKXNTLQKRVLVTVLEFL---PLLP-----SGILWLRLKIQSENVNATIYPG- 59
Qy 321 RDHMEALRSFLDPGSGYSWQ----DAHADVGHVLVGTGLGRVTECLSLDKLEAAPSEAAV 376
Db 60 -QSIQELPLFTFPFPFGDTWELAVIPSHSAAKTVTETVRALVINM-RVRGPFSEKDFE 117
Qy 377 SRALQLLAERFWAGVWFLGPDESSDTEHTPTDLPFGHVRIKIRMDIDVTR----- 429
Db 118 DYIRYDNCSSSVLAAVFHFPHNHS---KEPLPLAVKYHLRFSTRRNYMTOTGSPFLK 174
Qy 430 -----TNKIRDRWDPCPADPLT---DLRYVMGCVFVYLQDLVERAAVRLSGAN--- 476
Db 175 ETEGWHTTSLPPLFPNFGP-REPTSPDGGEGYIREGFLAVOHAVDRAIMEYHADAATRO 233
Qy 477 --PRAGLYLOOMPYPYCVDDVFLRVLRSPLPLFTLAWIYSVTLTVKAVVREKETRLRDT 534
Db 234 LFQRLTWTIKRFPYPPTIEDPFLVAIOVQLPLLLLSFTYTALTARAAVQOEKERRLEY 293
Qy 535 MRANGLRAVLWLGWFLSCIGLPFLLSAALLVLVLKLG-----DLTPYSHGCVFLFAAF 589
Db 294 MRMMGLSWLHWSAWFLFFFLFLIAASFMTLLFCVKVKPNVAVLSRDSFLVLAFLCF 353
Qy 590 AVATVTSQFLSAPFSRANLAAACGGLAYFSLYPYVLCVAVWRDLRPAAGRVAASLLSPV 649
Db 354 AISTISFVSTFFSKANMAAAGGFLYFTYIPYFVAPRYNMWTLTSLQKLCCLLSNV 413

Qy 650 AFGFGCESLALLBEQCEGAQWNVGTRPTA-DVFSLAQVSGILLDDAALYGLATWYLEAV 708
Db 414 AMAMGAQLIGKFEAKGMIQWRDLLSPVNVDDDFCFQGVLMGLLSVLYGLVTWYMEAV 473
Qy 709 CPGQYGIPEPNWPFRRSYWCWGRPPKSPAPCPTPLDPKVLV-----EEAPPGLSPGVSV 763
Db 474 FPGQFGVPQWYFFIMPSTWCG-KPRAVACKEEEDSDPEKALRNEYFEAPPEDLVAGIKI 532
Qy 764 RSLEK--RFEPSQPALRGLSLDFYQGHITAF LGHNGAGKTTTSLSGLSGLFPSPGGSFAFI 821
Db 533 KHLKSVFRVGNKDRAAVRDLNLNLNLYEGQITVLLGHNGAGKTTTSLMLTGLFPPTSGRAYI 592
Qy 822 LGHDVRSMAAIPRHLGVCPQYVNLFDMLTVDEHWFYGRKLGSLAAVWVGPQDRLLQDV 881
Db 593 SGYEISQDMVQIRKSLGLCFQHDILFDNLTVAEHLFYAQLKGLSRQKCEPEVKQMLHII 652
Qy 882 GLVSKOSVOTRHLSSGGMORKLSVAIAFVGGSOVVIDEPTAGVDPASRRGIWELLKYRE 941
Db 653 GLEDKWSRSRFLUSGGMRRLSIGIALIAGSKVLIIDEPTSGMDAISRRAINWLLQROKS 712
Qy 942 GRTLILSTHLDRAELIGDRVAVVAGGRLCCGSPFLRRHLSGGYYTLTLVKARLPITTN 1001
Db 713 DRTIVLTTHFMDRADLLGDRIAIMAKGELQCCGSSFLFKQYKAGYHMTLVKE----- 765
Qy 1002 EKADTMEGSDTRQEKONGSQSRVGTPTOLLALVOHWVPGARLVBEELPHELVLVLPYTG 1061
Db 766 -----PHCNPEDISQLVHHVNPATLESSAGAELSFILPRES 802
Qy 1062 AHDGSPATLPRELDTLRLAELRLGYSIDTSLREIFLKV----- 1100
Db 803 TH--RFEGLFAKLEKKQKELGFIASFGASITTEVEFLRVGKLVDSMDIOAIQLPALQYQ 860
Qy 1101 VEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETALENGEPAGSPAPETDQSGSPD 1160
Db 861 HERRASDWAVDSNLG-----AMDPSDGIGALIEBER---T 893
Qy 1161 AVGRVQWALTROOLOALLKRFLLARRSRRLGFAQIVLPALFVGLIAVPSLVPPFGHY 1220
Db 894 AVKLNTGLALHCOQFWAMFLKKAAYSWREWKVAAQVPLTCTVTLALLAINYSSELFDD 953
Qy 1221 PALRLSPTMYGAOVSPFSEDAPGDGRRARLLEALQEALEEPVVOHSSHRFSAPEVPAE 1280
Db 954 PMLRLTIGEYGRVTVFVSV-----PGTSQLGQOL-----SEH----- 985
Qy 1281 VAKVLASGNWTPSPSPACOCQPGARRLLPDCPAAAGPPPPPOAVTSGSEVVQNLTRN 1340
Db 986 -----LKDALQABEG--QEPREVILG-----D 1003
Qy 1341 LSDPLVKTYPRVROGLTKTKWVNEVRYGGFSLGGRDPGLPQSGELGRSVEELWALLSPL 1400
Db 1004 LEBFLI-----FRASVEGGGFN----- 1020
Qy 1401 PGCALDRVLKNTAWAHSLDAODSLKTFNKNKGWMAFVFNRAASNAIRLAHLPPGPARH 1460
Db 1021 -----ERCL--VAASFRDVCERTVVALFNQAYHSPATALAVVDNLLFK--LLCGP--H 1069
Qy 1461 AHSITTLNHP-----LNLTKESLEAALMASSVDVLVSIQVWFAMSVFSPASFTLVLEER 1515
Db 1070 A-SIVVSNFPOPRSALQAQAKDQNEG--RKGFIDIALNL--LFAMAPLASTFSILAVSER 1123
Qy 1516 VTRAKHLQLMGGLSPTLYWLGNTFMDMCNVLVPAACIVVLIFLAQOQRAYVAPANLPALL 1575
Db 1124 AVQAKHVQVSGVHVHSAFSLSALLMDLISFLISLLLVVFKAFDVRFAFRDGHMATLL 1183
Qy 1576 LLLYXGWSITPLMPYSPAFFSPSTAYVLTCTNLNLFINGINGSMATFVLEL---FSDOKLQ 1632
Db 1184 LLLYXGWAIIPLMYLNMFFLGAATAYTRLTIFNLISGI-----ATFLMVTIMRIPAVKLE 1239
Qy 1633 EVSRILKQVFLIIPPHFCLGRGLDMVRN-----QAMADAFERLGDQFQSP-LRWE 1682
Db 1240 ELKSLTDHVLFLPNHCLGMAVSSFVENYETRRYCTSSSEVAAHYCKKNIQYQENFYAWS 1299

Db 415 RDLSPVNVDDFCFGQVGLMLLDLSVLYGLVTWYMEAVFPQGVGPQWYFFIMPSYMC 474
QY 730 GPRPKSPAPCTPLDPKVLV-----EEAPPGLSGVSRSLEK--RFGPSQPALRGUS 782
Db 475 G-KPRAVAKKEEDSDPEKALRNEFEAPEDLVAGIKIKHLSKVRVGNKDRAAVRDLN 533
QY 783 LDFYQGHITAFIAGHNGAGKTTTSLSGILFPPSGGSFAFILGHDVRSSMAAIRPHLGVCPC 842
Db 534 LNLVEGQITVLLGHNGAGKTTTSLMLTGLFPPTSGRAYISGYEISQDMQVIRKSLGLCPQ 593
QY 843 YNVLFDMLTVDHNVWFYGLKGLSAVVCPEODRLLQDVGLVSKSVQTRHLSGGMQKRL 902
Db 594 HDILFDNLVABHLYFAQLGRKQKCEEVYQMLHIIIGLEDKWNRSRFLSGGVRRLK 653
QY 903 SVIAFAVGSSQVVIDEPTAGVDPASRRGIWELLKRYREGRTLILSTHLLDEABELLGDRV 962
Db 654 SIGIALIAGSKVLIDEPSTGMDAISRAIWDLLQKQSDRIVILVTHFMDADLLGDRI 713
QY 963 AVVAGRLCCGSPILFRLRHGSGYYLTLVKARLPLTTNEKADTDMEGSVDTREQKNGS 1022
Db 714 AIMAKGELQCCGSSLFLKQKYGAGYHMTLVKE-----745
QY 1023 QGSRVGTQQLLVALVQHWYPCARLVELPHELVLVLPYTCAGHDGSPATLPRELDTLAEIR 1082
Db 746 --PHCNPEDISQLVHHVHPNATLESSAGAELSFILPRESTH--RFGELPAKLEKQKELG 801
QY 1083 LTGYGSDTSLSEIFLKV-----VEBCAADTMDGDCGSOHLCT 1121
Db 802 IASFASITTMEEVFLRVGLKLVDSMDTQAIQLPALQYQHERRASDVAVDSNLCG-----856
QY 1122 GIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGPDAVGRVQGWALTROQLQALLK 1181
Db 857 -----AMDPDSDGIGALIEER---TAVKLTGLALHCCQFWAMFLK 894
QY 1182 RELLARRSRGLFAQIVLPALFVGLALVPSLIVPFFGHYPALRLSPTHYGAQVSPFSEDA 1241
Db 895 KAAYSWREKWKMAVQAQVPLTCTVLTALLAINYSSELFDPMRLRLTLGEYRTVVPFSV--952
QY 1242 PGDQGRARLLEALLQEALEPPVQHSHRFSAPVPAEVAKVASGNWTPESPSPACQC 1301
Db 953 -----PQTSQLCQQL-----SEH-----965
QY 1302 SQPGARRLLPDCPAAAGPPPPQAVTGSGEVVQNLTRNLSDPLVKTYPRLVROGLKTKK 1361
Db 966 -----LKDALQAEQ--QEPREVLG-----DLEFFLI-----989
QY 1362 WNEVRYGFSGLGRDPLGQSGELGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDA 1421
Db 990 FRASVEGGGFN-----ERCL--VAASFRDVG 1014
QY 1422 QDSLKIWFNNKGMHSMVAFVNEASNAILRAHLPDGPBARHHSITTLNHP-----LNLTK 1476
Db 1015 RTVNALFNNQAHSPATALAVVDNLLFK--LDCGP--HA-SIVSNFPQPSALQAARD 1069
QY 1477 QLSAALMASSVDLVISCVFAMSPFVPLVLIBERTVTRAKHLQMLGGLSPLTYMLG 1536
Db 1070 QFNEG--RKGFDIALN--LFAMAFLASTFSLAVSERAQKHQVQSVGVHVASFWLS 1124
QY 1537 NFLDMCNVLPACIVLIFLAFQORAYVAPANLPALLLLLLLYGWSITPLMYPASPFPS 1596
Db 1125 ALLMDLISFLIPSLLLVVFKAFDVRATFDGHMADTLLLLLYGWAIIPLMYLNNFFFL 1184
QY 1597 VPSTAVVLTCLNLFINGSMATVLEL---FSDQKQEVSRILKQVFLFPFHPFCLGRG 1653
Db 1185 GAATAYTRITIFNLSGI-----ATFLMTIMRIPAVKLEELSKTLDHVLPLPNCLGMA 1240
QY 1654 LIDMVRN-----QAMADAFLRGDRQFQSP-LRWEV--VGKNLLAMVITQ-----P 1696
Db 1241 VSSFYENYETRYCTSSSEVAHYCKKYNIOENFYAMSGVGRFVASMASGCAYLIL 1300
QY 1697 LFLFTLLLOHRSQLLPQVRV-----SLPLGEEDEVDARERVRVQATQG---D 1745
Db 1301 LFLIETNLLQRLGILCALRRRTLTLYTRMPVL-PEDQDVADERTRILAPSPDSLLHT 1359

RESULT 13

US-08-762-500-25
; Sequence 25, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/762,500
; FILING DATE: 17-JUN-1996
; APPLICATION NUMBER: PCT/US96/10469
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1684 amino acids
; TYPE: amino acid


```
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-8

Query Match      3.9%; Score 427; DB 4; Length 328;
Best Local Similarity 34.9%; Pred. No. 3.2e-29;
Matches 112; Conservative 59; Mismatches 142; Indels 8; Gaps 5;

QY 1750 RNLTQVYRGQRMVADRLCLGTPPECFCGLLVGNAGKSTSTFRMTVGTDLASRGEAVLAG 1809
DB 12 RDLVVFVGQTR--AVDGLDLVVRAGTIHGVLGPNAGAGKTTAIKMLATLMRPTSGTASVLG 69
QY 1810 HSAVAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARLGVPEAQVAQTAGSGLARLGL 1869
DB 70 HDVREAAEVRRIIGLTQGTMSVEDMTGVQNLILAGRLQGLRHASAAARAEQLMEAFDL 129
QY 1870 SWYADRPAGTYSGGNKRKLATLALVGPVAVVFLDEPTTGMDDPSARRFLWNSLLAVVREG 1929
DB 130 TEVGRVLVKTFTGGGRRRIDVAASVVTPELFLDEPTTGLDPRSRSEWEMIRALVRDG 189
QY 1930 RSVMLTSHMECEALCSRLAIMVNGFRCLGSPQHLKGRFAAG-HTLTLRVPAARSQPA 1988
DB 190 GTVLLTTQVDEADHDLABELTLIDHGRIVAQGTTPPELKASRAAGVLDVLRDPERRADAG 249
QY 1989 AAFVAAEPGSELREAHGRLRFQLPGRCALARVFGELAVHGAHGVDESVSOTMLE 2048
DB 250 ALLAKAVGAAADL-DSDPARLSVRVTDPPRAALA--LGELARAGTH--VDDFTLQPSLD 304
QY 2049 EVFLYFSKQDQKDETEOKE 2069
DB 305 TVFLALTGHSTVDASEEBAE 325

RESULT 15
US-09-252-991A-18351
; Sequence 18351, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18351
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18351

Query Match      3.8%; Score 414; DB 4; Length 607;
Best Local Similarity 17.7%; Pred. No. 1.4e-27;
Matches 223; Conservative 100; Mismatches 235; Indels 702; Gaps 26;

QY 732 RPPKSPAP-----CTPLDPLKVLVEEAPPGLSPGVSVRSLEKRFPGSPQPALRGLSLDFY 786
DB 3 RSMRPAPVRSRSGVEMQVSPAIAVPAAGADEAAVIEDVKHF--GDVKALRGLRSARIH 60
QY 787 QCHITAFHGNAGKAGTTLTSLSLGFLPPSPGGSFATLGHVDRSSMAAIRPHLGVCPOYNVL 846
```

```
DB 61 YGRLTGLVGPDCGAGKTLMRILITGLVNPAGRVTLTLAGYDVVKDNDAIHVASGVMQRFGL 120
QY 847 FDMLTVDHWHVYFGRKLGLSA-----AVVGEQDRLLQDVLGSKQSVQTRH 893
DB 121 YEDLSVMENMRLYAQURGMDADRNAELFAELLDFTRLGPFTRKL-----AGK 167
QY 894 LSGGMQRKLSVAIAFWGGSQVVILDEPTAGVDVPASRRGIWELLKYR-EGRTRLILSTHHL 952
DB 168 LSGGMKQKLGACALMARPKVLLDDEPGVVDVPVSRQDLRWVQALTDGMAVVMSTAYL 227
QY 953 DEAEELLGDRVAVVAGGRLCCCGSPFLRRRLHSGSYLTLVKARLPLTTNEKADTDMEGSV 1012
DB 228 DEAEER-----CESVLLNQ----- 241
QY 1013 DTRQEKKGSGQSRVGTTPQLLALVQHVWPGARLVEELPHELVLVLPYTCAGHSGSFATLPR 1072
DB 242 ----- 241
QY 1073 ELDTRLAELRLTGYGISTSLLEEIPLKVVEECAADTDMEDSCGQHLCTGIAGLDVTLRL 1132
DB 242 -----GQLLPDG----- 248
QY 1133 KMPQETALENGEPAGSAPETDQGGSPDAVGRVQGWALTROQLQALLKRFLLARRSRG 1192
DB 249 --PPQE-----LTAQ-----LEGRS--- 261
QY 1193 LPAQIVLPALFVGLALVFSLVPPPGHYPALRLSLTMYGAQVSPFSEDAPOGPGARLLE 1252
DB 262 -----FLENVGAERRAVLTE 277
QY 1253 ALLQEALEEPPVQHSRFSAPAEVPAEVAKVLASGNWTPESPSPACQCSQCGARRLLPD 1312
DB 278 AL-----D 280
QY 1313 CPAAAGSGPPPPQAVTSGSEVQNLTRNLSDFLVKTYPLRVQGLKTKKWNVRYGGSF 1372
DB 281 LPSVSDG-----VIQAGVRV-----VLREGAPTEQ----- 306
QY 1373 LGGRDPGLPSQELGRSVEELWALLSLPLPGGALDRVLKNTAWAHSLDQDSLKTWPNK 1432
DB 307 ----- 306
QY 1433 GWHSMVAFVNRASNAILRAHLPPGPARHAHSITTLNHLPLNLTKEQSEAAUMASSVDVLV 1492
DB 307 ----IQALADRA-----QVQLAPVDPAR----- 324
QY 1493 SICVVFAMSFVPASFTVLVIEERVTRAKHLQMGGLSPTLYWLNFLWDMCNLYLPACTIV 1552
DB 325 -----FEDAFI-----DLGG----- 335
QY 1553 VLIFLAFQORAVVAPANLPALLLLLLLLYGSITPLMYPASPFVSPSTAYVVLTCINLPI 1612
DB 336 ----- 335
QY 1613 GINGSMATFVLELFDQKLQBSVRLKQVFLIFPHFCLGRGLDMVRNQAMADAPERLGD 1672
DB 336 ---GPGGTSLA-----ERL--- 347
QY 1673 RQFOSPLRWEVVGKNLLAMVIOGPLFLTLTLLQHRSQLLPQPRVRSPLLGEEDDVAR 1732
DB 348 ---SPV-----ELGSDVA- 357
QY 1733 ERERVVQATOGDVLVLRNLTKVYRGQRMVADRLCLGIPPGECGLLVGNAGKSTSTFR 1792
DB 358 -----VSCNLTK--RFGFTATDQVSFEVQKGEIFGLLGPNGAGKSTTFK 401
QY 1793 MVTGDTLASGEAVLAGHSVAREPSAAHLSMGVCPQSDAIPELLTGREHLELLARLGV 1852
DB 402 MLCGLLKFTAGEAHVGHDLRHATGAQSKQCYWAQKFSLYGLLSVQRNLEFSAGVYGLE 461
QY 1853 EAQVAQTAGSGLARLGLSWYADRPAGTYSGGNKRKLATLALVGPVAVVFLDEPTTGMDDP 1912
```

Db 462 GNVRRERIEEMIAFDLGDWLSATPDSLPLGHKQRLALACSLMHRPPVLFDEFTSGVDP 521
Qy 1913 SARFRLNLSLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCIGSPQHUKGRFAA 1972
Db 522 ITREFWTHINGLARKGVITVTHFMDAE-YCDRVAMLSRARLIALDTPDALK-RIAA 579

Search completed: December 29, 2004, 22:54:05
Job time : 43.2021 secs

354 CMWFEI SPQIWTWEMNSQEMDI YB TTI LNSPBNDOCEWEQVI DCI DWTACDITM B PT AKNPEB 413

```
Qy 332 --DPGSGYQWQADHADVGLVGLGRVTECLSLDKLEAAPSEAAALVSRALQLLAEHFW 389
Db 414 VQSPNGSVYTWREAFNETQAIQITISRFMECVNLKLEPIETVRLINKSMELLDERKFW 473
Qy 390 AGVVFLLPEDSSDPTHTPDLG--PGHVIKIRMDIDVTRTKIRDRFWDGCPAADPL 447
Db 474 AGIVFTG-----ITPDSVELPHVVKYKIRMDIDNVERTNKIKDGYWDGCPRADPF 523
Qy 448 TDURYVMGFGVYLQDLVERAAVRVLGANPRAGIYLQOMPYPCVVDVDFURVLRSPLPF 507
Db 524 EDMRYVMGFGAYLQDVVEQAIIRVLTSSEKTKGVYVQOMPYPCVVDVDFURVMSRMPLF 583
Qy 508 LTLAWIYSVTLTKVAVREKETRLDRMTWRANGLSRAVLWGLFSLCLGPELSSAALLVLV 567
Db 584 MTLAWIYSVAVIKSIYVEKEARUKETMRIMGLDNGILWFSVSSLIPLLVSAGLLVI 643
Qy 568 LKLGDIIPYHPGVVFLFAFAVATVTSQFLLSAPFSRANLAAACGLAYFSLYLPYVL 627
Db 644 LKGNLIPYDPSPVVFVFLSVFAMVTILQCFLLSTLFSRANLAAACGIIYFLYLPYVL 703
Qy 628 CVAWRDLRPAGRVAASLLSPVAFPGCESLALLEGEQGAQMNVTGRT-ADVFSLAQ 686
Db 704 CVAWQYVGFSEIKIFASLLSPVAFPGCEYFALFEEQIGVQMDNLFESPVEEDGNLTT 763
Qy 687 VSGLLLDAAALYGLATWYLEAVCPGOYGIPEPWNFPFRSVMCWPCKSPACPCTPLD- 745
Db 764 AVSMGLDFTLYGVMTWYIEAVFPQOYGIIPRPYFPCTKSYFWGEEIDEKSHPGSSOKGV 823
Qy 746 PKVLVEEAPPGLSPGVSRSLERKFPGPSOPALRGLSLDFYQGHITAFLGHNGAGKTTTL 805
Db 824 SEICMEEPHLRLGVSIQNLVYKVRDGMKVAVDGLALNFVEGQITSFLGHNGAGKTTM 883
Qy 806 SILSGLPPPGSGSFIILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTVDENHFWFYGRUKGL 865
Db 884 SILTGLFPPTSGTAYILGKDIRSEMSIRQNLGVCPQHNVLFDMLTVEEHIWFYARUKGL 943
Qy 866 SAAVGEQDRLLQDVLG-VSKQSVQTRHLSSGNORKLSVAIAFVGSQVVLDEPTAGV 924
Db 944 SEKHVAEMEQMALDVGLPPLSKLSKTSQSLSGGMORKLSVALAFVGSQVVLDEPTAGV 1003
Qy 925 DPASRRGIWELLKYRGRTRILLSTHDLBAELLDGKVAVVAGGRCLCCGSPFLRRHLG 984
Db 1004 DPTSRRIWELLKYRGRTRILLSTHDEADILGDRALIIISHGLCCVGSUFLKNQGL 1063
Qy 985 SGYYTLVLKARLPITTN-----EKADTMEGSSVDTRQEKKNKGSGSRVGPQL 1032
Db 1064 TGYTLVLKXDVESLSRNSSTVSCLKEDSVSQSSDAGLGDSDHSDTLTIDVSAI 1123
Qy 1033 LALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATILPRELDRLAELRLTGIGISDTs 1092
Db 1124 SNLRKHVSEARLVEDIGHELYVLPYEAKEGAFVELFHEIDRLSDLGISYIGISETT 1183
Qy 1093 LEEIFLKVBECAADTOWEDQ-----SCGOHLCTGIAGLDVTLRLKMPQ 1137
Db 1184 LEEIFLKVBEESGVDAETSDGTLPARNRARFGDKQSC-LHPFTEDDAVD-----PN 1234
Qy 1138 ETALENGEPAGSAPETDQGGPDVAG--RVQGWALTQQQLQALLKFLRLAARSRRGLFA 1195
Db 1235 DSDID-----PESRETDLGMDGKGYQLGKWLTKQQFVALLWKELLITARRSRKGFFA 1289
Qy 1196 QIVLPALFVGLAVFSLIIVPPGHYPALRLSPTHYGAQVFFSDEADGDPGRARLLEALL 1255
Db 1290 QIVLPVAVFCIALVFSLIIVPPGHYSLELOPMWYNEQYTFVSNADPEDMGTOQLNALT 1349
Qy 1256 QEAG-----LEPPPVQSHSRFSAPVPAEVAKVASGNWTPESPSPACQSQPG 1305
Db 1350 KDPFGFTRCMEGNIPDTPCLAGEEDWTISPVQSIIVDLFQNGWNTWKNFSPACQCSDK 1409
Qy 1306 ARLLLPDCAAGAPPFPQAVTGSGEVQNLGTGRNLSDFLVKTYPRILRVGLKTKKWNE 1365
Db 1410 IKKMLVPCPPGAGGLPFPQQRKQTADILQNLTGRTNISDYLVTYVQIIAKSLKNKIWNE 1469
```

```
Qy 1366 VRYGFGSLG-GRDPLGSLGQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLDQDS 1424
Db 1470 FRGFGSLGVSNQALPPPSHEVNDAIKQMKLLKLTWDTADRFSLSLGRFMAGLDTKNN 1529
Qy 1425 LKIWFNNKGWHSVAVFVNRASNAIILRAHLPPGPARHAHSITTLNHPNLNLTKBQLSEALM 1484
Db 1530 VKWFNNKGWHAISFLNINNAIILANLQKGNPSQYGITAFNPLNLTKQQLSEVALM 1589
Qy 1485 ASSVDVLVSI CVVPAMSFVPASFTLVLIERVTRAKHLQLMGGLSPTLYLWGLHFLWMCN 1544
Db 1590 TTSVDVLVSI CVTFAMSFVPASFTLVLIERVTRAKHLQLMGGLSPTLYLWGLHFLWMCN 1649
Qy 1545 YLVPACTVLVILFAFOORAVVAPANLALLLLLWGSITPLMYPASFFSFPSTAYVV 1604
Db 1650 YVVPATLVIIIFCFQOKSVSVSTNLPVALLLLWGSITPLMYPASFFVKLPSTAYVV 1709
Qy 1605 LTCINLFIGINGSMATFVLELSDQKLOEYSRLKQVFLIPPHFCLGRGIDMVNRNQMA 1664
Db 1710 LTSVNLFIGINGSVATFVLELFTNNKLNINDILKSVFLIPPHFCLGRGLIDMVNRNQMA 1769
Qy 1665 DATERLGRDROFQSPLEWVGVKNLLAMVIOGPLFLLFTLLLOHRSQLLPQPRVRSPLIG 1724
Db 1770 DALERFGENFVSPFLSMDLVGRNLFAMAVEGVVFFLTIVLIQVRFIRPRPVKAKLPLN 1829
Qy 1725 EEDVARERERVQOGATQGDVLVRLNLTKVYRGQRMPAVDRLCLGIPPGCEFGLLGVNG 1784
Db 1830 DEDEDVRRERQRLIDGGQNDILEIKELTKLYRKRKPAVDRLCIGIPPGCEFGLLGVNG 1889
Qy 1785 AGKTSTFRMVGTDLASRGEAVLAGHSVAREPSSAAHLSMGYCPQSDAIFELLTGREHLEL 1844
Db 1890 AGKSTTFKMLTGTPTVTRGDAFLNKNLSILNIEHVQNMGYCPOFDAITELLTGREHVEF 1949
Qy 1845 LARLGVPEAOVQAQTAGSGLARLGLSWYADRPAGTYSGNKRLATLALVGDPAVVFLD 1904
Db 1950 FALLRGVPEKEVGKFGESAIKRLGLVKYGEKYSNYSNGNKRKLSTAMALIGGPPVVFLD 2009
Qy 1905 EPTTGMDPSARRFLWNSLLAVVREGSVMLTSTSMEECEALCSRLAMVNGRFRCLGSPQ 1964
Db 2010 EPTTGMDPKARRFLWNCALSIVKEGSRVLTSTSMEECEALCTFMALVNGRFRCLGSPQ 2069
Qy 1965 HLKGRFAAGHTLRLVPAARS--OPAAAFVAAFPFGSELBEAHGGRURLFPFGRCRICALA 2022
Db 2070 HLKNRFGDGYTVIVRIAGSNPDLKPVQEFFGLAPFGPSVLKKEKRNMLQYQL-FSSLSLA 2128
Qy 2023 RVFGLAVHGAEGVEDFVSQTMLEEVLYFSKQCKDE---DTEQKEAGVGVDPAPG 2079
Db 2129 RIFSILSQSKRRLHIEDYSVSQTLQGVFNFAKQSDDDHLKDLSLHKQTV-VDVAV- 2186
Qy 2080 LQHPKRVSQFLDDPSTAET 2098
Db 2187 -----LTSFLQDEKVKES 2199
RESULT 2
A59189
ATP-binding cassette transporter - human (fragment)
N;Alternate names: KIAA1062 protein
C;Species: Homo sapiens (man)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C;Accession: A59189
R;Kikuno, R.; Nagase, T.; Ishikawa, K.; Hirose, M.; Miyajima, N.; Tanaka, A.; Kotani,
DNA Res. 6, 197-205, 1999
A;Title: Prediction of the coding sequences of unidentified human genes. XIV. The complete
A;Reference number: Z22961; MUID:99397452; PMID:10470851
A;Accession: A59189
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1529 -KIK-
A;Cross-references: UNIPROT:Q9HC28; GB:AB028985; NID:95689460; PIDN:BA83014.1; PID:d104
A;Experimental source: chromosome 9; clone hJ03579; clone lib pBluescriptII SK plus; tis
C;Genetics:
A;Map position: 9
A;Note: KIAA1062
```

Db 900 PASFVFLVAEKSTKAKHLQFVSGCNPDIYWLANYVMDMLNVLVPATCCVILFVFDLPA 959

Qy	261	LMAQVNTFEBELTLLRDVREWEMLGPRIFTFMNDSSNVAMLQRLLOQWDEGRQPRPGG	320
Db	9	LLLWKNTYLOKRVLVTVLELFL-----PLLP-----SGLLWLRUKIOSENVNPNATYIPG-	59
Qy	321	RDMNEALRSFLDPGSGGYSWO----DAHADVGLHVGLTRVTECLSDKLEAPSAALV	376
Db	60	-QSIOELPLPTFPFGPDTHELAVIPSHSDAAKTVTETVRALVINM-RVRGPFSEKDFE	117
Qy	377	SRALQLLAHRFWAGVVFGLPEDSDPTEHPTDLPGLGPHVRIKIRMDIDVVVTR-----	429
Db	118	DYIRYDNCSSVLAADVFEHPFNHS---KEPLPLAVKYHLRFSYTRRNVMWTQGSFPLK	174
Qy	430	-----TNKTRDRFWDGPAADPLT-----DLRYWGGFVYLODLVERAAVRVLSCAN---	476
Db	175	ETEGWHTTSUPLPLPNFGP-REPTSPDGBEGYIREGFLAVQHVAHDAIRMEYHADAA	233
Qy	477	--PRAGLYLOOMPYPCCVDDVFLRVLRSPLFLTALWIYSVTUTKAVVREKETRLRDT	534
Db	234	LFQRLVTIKRFPYPPIADPFLVAIOYQLPLLLLLFTYTALTIAAAVQEKERRLKEY	293
Qy	535	MRAMGLRAVLWLQWFLSCIGPFFLLSAAALLVLVLKLG-----DILPYSHPGVVF	589
Db	294	MRMGLSSWLHWSAWFLFFLFLLLIAAASFMTLLFCVKVKPNVAVLSRSDPSLVLAFL	353
Qy	590	AVATVOTOSFLLSAPFSANLAAACGLAYSLYLPLYVLCVNRDRLPAGGRVVAASILSPV	649
Db	354	AISTISFWMVSTPFSKANMAAAGGFLYFTYIPYFFVAPRYNMWTLQKLCSCLLSNV	413
Qy	650	AFGFGCSLALLBQBGCAWHNVGTRPTA-DVPSLAQVSGLLILLDAALYGLATWYLEAV	708
Db	414	AMANGAQLIKGFEAKGNGI QWRDLLSPVNVDDEFCFGVQLGMLLSDSVLXGLVTWY	473
Qy	709	CPGOYGIPEPNPFPFRASYCGPRPPKSPACPTPLDPKVLV-----EEAPPLSGPVSV	763
Db	474	PPGOYGVQPWYFIMPYSYWC-KPRAVAGKEEDSDPEKALRNEYEAPSPDLVAGIKI	532
Qy	764	RSLEK--RFPGSPQALRGILSDPYOQHITAFIGHNGAGKTTTLLSGLPSPSGSAFI	821
Db	533	KHLSKVFRVGNKRAAVRDLNLIYEGQITVLLGHNGAGKTTTLLSMUTGLPPTPSGRAYI	592
Qy	822	LGHDRSSMAAIRPHLGVCQYNNVLFOMLTVDEHWFVYGRLLKGLSAVAVGPEODRLIQDV	881
Db	593	SGYEISQDMVQIRKSLGLCQHDILFDNLTVAEHLHYAQLKGLSRQKPEEVKQMLHII	652
Qy	882	GLVSKQSVQTRHLSGGWOKLSVAIAFVGSGQVILDEPTAGVDPAARRGIWELLKYRE	941
Db	653	GLEDKWNRSRFLSGGWRRLKLSIGIALIAGSKVILDEPTSGMDAISRRAIWDLLOKRS	712
Qy	942	GRTLILSTHLDRAELLDGAVAVVAGRLCCGSPFLRRLHSGSYVLTAVKARLPLTTN	1001
Db	713	DRITVLTHFMDAHLGDRIAMAKELQCCGSLFKOKYGGYHMTLVKE-----	765
Qy	1002	EKADTDMEGSVDTREQKNKSGQSRVGTPLLALVQHWVPICARLVEBPHLVLVLYPTG	1061
Db	766	-----PHCNPEDISQLVHHVFNPNATLESSAGAELSFILPRES	802
Qy	1062	AHDSGPAFLPRELDRLAELRLTYGIGSDTSLEBIFLKV-----	1100
Db	803	TH--RFBGLFAKLEKKELGIASFSGASIITMBEVFLRVGLVDSMDIOAIQLPALQYQ	860
Qy	1101	VEECAADTDMEDGCGQHLCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETDQSGPD	1160
Db	861	HERRASDWA VDSNLCG-----PHCNPEDISQLVHHVFNPNATLESSAGAELSFILPRES	802
Qy	1161	AVGRVQGWALTROQLOALLKRFLLARRSRRLGPAQIVLPALFVGLAVFSLVPPFGHY	1220
Db	894	AVKLNTGLALHCQQFWAMELKKAAYSRWKMWAAQVLPVLTCTVLTALLAINYSSEL	953
Qy	1221	PALRLSPMTYGAQVFSFSEADPGDGRARLLEALLQEAAGLEBPPVQHSRRFSAPEVPAE	1280
Db	954	PMRLTLTGXYGRTVVPFSV-----PGYSOLGQQL-----SEH-----	985

Qy	1281	VAKVLASGNWT	PPSPSPACOCSPQ	CARRILLPDCP	AAAGGPPPPQAVT	GSVEQVQNL	TGRN	1344																					
Db	986	-----	-----	-----	-----	-----	-----	1003																					
Qy	1341	LSDFLVKTY	PRLVRQGLKTK	KWNVRYGFS	LSGRDGLPSG	QELGRSVEEL	WALLSPL	1400																					
Db	1004	LEEFLL	-----	-----	FRASVEGG	GFN	-----	1020																					
Qy	1401	PGGALDRVL	KNLTAWAHS	LDAQSLKTIW	FNNKGWHS	WVAFVNR	ASNAILLRAHLP	PPGPARH	1460																				
Db	1021	-----	ERCL	--VAAS	FRDVGERTV	WVALFNQAYH	SPATALA	VVDNLLFK--LLCGP--H	1069																				
Qy	1461	AHSITTLN	HP-----	LNL	TKQLSEAL	MASSVDVL	VSICVVFAMS	VPVPSFTVL	VLEER	1515																			
Db	1070	A-SIVSNF	POPRA	QAQKQNEG	---RKG	FDIALNL--	LFAM	PLASTS	FLVLSR	1123																			
Qy	1516	VTRAKHQL	QMGGLSPT	LYMIGNFL	MDMNYLV	PACIVVL	FLAFQORAY	VYAPANL	PALLL	1575																			
Db	1124	AVQAKHV	QVSGVHV	ASFWSALL	MDLISPL	SFLLSLL	VVVKAFD	VRAFTRD	GHMADTLL	1183																			
Qy	1576	LLLLYGS	ITPLMY	PASFPF	SPSTAYV	UTCINL	FIGINGS	MATFVLEL--	FSQOKLQ	1632																			
Db	1184	LLLLYGM	AIIPFL	MYLMNF	FLGAATAY	TRLTIF	NIILSGI	----ATFL	MVIMIR	PAVKLE	1239																		
Qy	1633	EVSRLKQ	VELPIPH	FCLGRGL	MDVRN	-----	QAMADA	FERLGD	RQFOSPL--RWE	1682																			
Db	1240	ELSKTDH	VLFLVPH	NHC	MAVSF	YENYTRY	CTSS	SEVAHY	CKKYNIOY	QENFVWS	1299																		
Qy	1683	V--VGN	LLAMV	IQG-----	PLFL	FTLLQHR	SQLLPQ	PRVR-----	SUPL	LGEBE	1727																		
Db	1300	APGVGR	FVSA	MAASG	CAYLL	LLFLIET	NLLQRL	GILCAL	RRRTLT	TELYTR	MPVL--PED	1358																	
Qy	1728	EDVARE	ERVQ	ATQG--	--DVL	VLRLN	TKVYR	QORP--	AVDR	CLCIGIP	PPGCFGL	LV	1782																
Db	1359	QVADER	TRTLAP	SPDS	LLTPI	IIELS	KVYE--	QRVPL	LA	VDRL	SLA	VQKGC	FLGLGF	1417															
Qy	1783	NGAKT	STPRM	VTGDT	LAS	GEAV	LASH	VSARE	PSAAH	LSMGV	CPQSD	DA	FELL	TGREHL	1842														
Db	1418	NGAKT	ITFM	LGT	ESR	LTSGD	FAVGGH	RIS	SDVK	VRQIRIG	YCPQ	FDAL	LDHMT	GREML	1477														
Qy	1843	ELLAR	LGVPE	QAQV	ATAG	SGLAR	LGLSW	YADR	PAGTY	SGGN	KRKL	ATAL	ALVGD	PAVVF	1902														
Db	1478	VMVAR	LGI	PER	HIG	ACVENT	LRGL	LLE	HANK	LVRTY	SGN	KRL	STG	LIGE	PAVIF	1537													
Qy	1903	LDEPT	GMD	SARR	FLW	NSLLA	VVREGR	SVML	TSHS	MEE	CEAL	CSRLA	IMVGR	FRCLGS	1962														
Db	1538	LDEP	STGMD	PVAR	RL	LWD	TVARE	S	GKAI	IT	TSHS	MEE	CEAL	CTRLA	IMVQ	QFCKLGS	1597												
Qy	1963	PQHLK	RFA	GA	CHT	L	TRV	PAAR	SQPA	-----	AA	VAAE	FP	GS	EL	REAH	GGRLR	PQLP	PPGR	2018									
Db	1598	PQHLK	S	KEG	G	YS	LR	AKV	QSE	Q	EALE	EEF	KA	F	VD	L	TF	GS	V	LE	DH	Q	GV	HYHL--PGD	1656				
Qy	2019	CALAR	V	GE	L	AV	H	AE	G	H	VE	D	F	S	Q	T	M	L	E	E	V	F	L	Y	F	S	2055		
Db	1657	LSWAK	FG	IL	E	K	E	K	E	K	Y	G	V	D	D	Y	S	O	I	S	L	E	Q	E	V	L	S	E	1693

RESULT 5

A59188
ATP-binding cassette transporter ABC3 - human
C:Species: Homo sapiens (man)
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change
C:Accession: A59188
R:Connors, T.D.; van Raay, T.J.; Petry, L.R.; Klinger, K.W.; Lan-
Genomics 39, 231-234, 1997
A:Title: The cloning of a human ABC gene (ABC3) mapping to chromo-
A:Reference number: A59188; MUID:97179225; PMID:9027511
A:Accession: A59188
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1704 <CON>
A:Cross-references: UNIPROT:Q99758; GB:U78735; NID:g1699037; PI
C:Genetics:

A:Description: The sequence of C. elegans cosmid Y39D8C.

A:Reference number: Z21408

A:Accession: T33783

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1802 <BEC>

A:Cross-references: UNIPROT:Q9TV8; EMBL:AF101313; PIDN:AAC69223.1; GSPDB:GN00023; CESP:

A:Experimental source: strain Bristol N2; clone Y39D8C

C:Genetics:

A:Gene: CESP:Y39D8C.1

A:Map position: 5

A:Introns: 45/3; 114/1; 195/1; 230/3; 543/3; 794/1; 849/1; 1036/2; 1099/1; 1132/3; 1165/

C:Superfamily: ATP-binding cassette homology

Query Match	16.8%; Score 1829; DB 2; Length 1802;
Best Local Similarity	28.4%; Pred. No. 4e-106;
Matches 539; Conservative 312; Mismatches 667; Indels 382; Gaps 55;	
QY	325 EALRSFLDPGGGYSWQDAHADVGHVLTGRVTECLSLD---KLEAAPSEAAALVS---R 378
DB	124 DATQIMDKIQNRYTANDLLNPAAVIGKLYNITAVPVLTNTNTYKGTTEGEMVSWMQG 183
QY	379 ALQLLAHRFWAGVVFGLPESSDTEPTDPLGCHVRIKRMIDIVVTRN-----KIR 434
DB	184 QFQSECDNPLLAGIVF-----DQSIADKLNPD-----KRDFTYIIRLSNTHRRSR 229
QY	435 DRFMDGPAADPLTD--LRYVMG-----GPLYQLDLVERAAVRVLSGA 475
DB	230 NAFGNSVPWTSVTSFAVQYVSGPINDDNGSGGYWQEGFMTVQRAVDVAITEITIGE 289
QY	476 N-----PRAGLY--LOOMPYPYVDDVFLRVLSRSLFLLTFLTAWISVTLTVKAVVREKETR 530
DB	290 DAQLTFLLDSYQVSRFPFGYSTKI--IEIGAFFMPVIVIFSPMSTVIVIVRAVVVEKEDR 348
QY	531 LRDTWRANGLRAVLWGLFSLGPFLLSAALVVLKGDILPYSPHGVPVFLAFAFA 590
DB	349 LKEYMRVNGLSQFINWVAHFI--INAYAKLTFAVITLIMHFVALKSDMTLMFVFLMTYA 406
QY	591 VATVTQSLLSAFFSRANLAAACGLAYFSVLYLPVVL--CVAWDRLPAGGRVAASLLSP 648
DB	407 FDVYFAFMISFMSATSATLISVFWMLLYFWAFSSIDQTNPYPLGYRLINCINPD 466
QY	649 VAFGFCBSLALLBEGQGAQHNVGTRPTAD---VFSLAQVSGLLLLDAALYGLATWYL 705
DB	467 IALNYGLQLLAAAYETAQADGLKXWGLFPTSPDNLLTFGHALIA--LIVDGIIMILTYI 524
QY	706 EAVCPQYGIPE--PWFPPFRSYWCGPRPKSPAPCPTPLD-----PKVLVEEAPPG 756
DB	525 EAVIPGGEGVPQKPMFFVL--PSYWF-----PNSGSKTVSDSQFOQIQIYADHVKLEKPTD 579
QY	757 LSPGVSVRSLEKRPFGS-----PQPALRGSLDFYOGHITAFIHLNGAGKT 802
DB	580 LIPTINVNLTKYGTSPFKFLDCKFKSGEKRAVSNLNLKMPYQOCTVILGHNGAGKS 639
QY	803 TTLSILGLFPFPGSGSAPFILGHDVRSSMAAIRPHLGVCPQYNVLFDMLTVDEHVMFYGLR 862
DB	640 TTFESMLTGAVSPSSGSAVYNDFDRTSLPKIRREMGCLCPQNTLFGFWVWEHLEFFAKL 699
QY	863 KGLSAAVVGPPQDR--LQDVGVLVSKQSVQTRHLSGGMORKLSVAIAFVGGSQVILDEPT 921
DB	700 K---ERTWDPBEAREILARLIDFKADFMAGALSGQKRLSLAIALTIGGSEVVMLEDEPT 756
QY	922 AGVDPASRRGWELLKYREGRTILSTHLDLAEILGDRVAVVAGGRLLCCGSPFLRR 981
DB	757 SGMPFGARHETWTLIQREKERTILLTTFHEEADLLGDRIAIMAHGQLECCGSPMFLKQ 816
QY	982 HLGSGYLLTVKARLPLTNTKADTDMEGSVDTTQEKNGSGQSRVGPFPQALLALVQHWPV 1041
DB	817 QYGDGYHLTV-----YDITSPDVSKTTD-----IIREYIP 848
QY	1042 CARLVEELPHELVLVLPVTGAHDGSGFATLRELDTRIALRLTGVIGSDTSLREIFLKV 1101
DB	849 EAHVFSYIGQBATYLLSAT--HRPITPKLFELEDHQTCCGITSFGVSIITMBEVFLKVG 906

QY	1102 EECADTDMEDGSCGHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGDPD- 1160
DB	907 HTADERYNYEHG-----IEN--DISEMIKDDPILQDL 937
QY	1161 -AVGRVQGMALTRQOQALLKRLARRRRRGLFAQIVLPALFVGLALVFLSIVPPFGH 1219
DB	938 RAQVRVTGFTLQMOHAKAMFYKRAIFFPFRKWTQPLQPLVFPVAVLVLMVFTSQVLP--- 993
QY	1220 YPALRLSPMYGAQVSPSESDAPGDGPRARLLEALLQEAAGLEEPVQHSRRSPAPEDPA 1279
DB	994 -----SVKE----- 997
QY	1280 EVAKVLASGNWTPESPSPACQCSOPGARRLLPDCPAAAGGPPPPQAVTGSGEVQNLTG- 1338
DB	998 -----QDQOTISLAFSDTKAG-----HLVSDSGNYVTLLGGS 1031
QY	1339 RNLSDFLVKTYPRL--VRQGL-----KTKKWVNEYVGGFSLGGRDPGLPSGOELGRSVEE 1392
DB	1032 QNLSSMWQGTVTQLGVTQTVVDITSNVEKFTMDQTN--AMGSRFTGLHYA----- 1079
QY	1393 LWALLSPULPGGALDRVLKNLTAWAHSLLDAQSLKIWFNNKWHSMVAFVNRASNAILRAH 1452
DB	1080 ----LGFVP-----SMFNFSVSVPSLKTFFNNFGLYTPALAITFTDSMILSQ- 1123
QY	1453 LPPGPARHAHSITTLNHP-----NLTKQOLSEAA--LMASSVDVLVSICVVPAMSFVP 1504
DB	1124 ----KQKQYQSFYAVNHPPLPSTQDTLKNVRSQGAALIAVGLIVSFAVCV----- 1171
QY	1505 ASPTLVLIIEERVTRAKHLQMLGSLPTLYLWLNFLMDCNLYLVACIIVVLFLAFQORAY 1564
DB	1172 AGYSQFLITERKKSKMQLLSGIRPMMFWLTAFIWDAAWFVIRILCFDAIFYIENITAY 1231
QY	1565 VAPNLPALLLL--LLLQWSITPLMPASFPSPSTAYVVLTCINLFIGINGSWATFVL 1623
DB	1232 THDFGVMILITLFLLYGWTALPFTYWFQFPFESAPKGFMMVTHILTMIGSIAVPII 1291
QY	1624 E-----LPSDQKLOEVSRI-----LKVFLFPHF-----CLG 1651
DB	1292 SQTSSLDAGYLWSIIFAWLFFTYNISQIATVTFQENVRIACKLDCTIPMKVATACG 1351
QY	1652 RGLDMVRNQAMADAFERL-----GDROFQSPLRWEVVGKNNLAMVI-----QGFLPL 1699
DB	1352 -----TASERLYVDNVLFVGNR-----KGIIVVYVFLAVQGFIVW 1386
QY	1700 LFTLLQLHR--SQLLPQPRVSL--PLLGEDEEDVARERE-----RVQOATQGD 1745
DB	1387 IWVFMRENDQFTKLIFALIRCKADNPINDITDITDKVDERVEDSDVIAEKSVVQRLANN 1446
QY	1746 --VLVLRNLTKVYRGQMPAVDRCLGIPGCECFGLGVNGAGKTSTFRMVTGDTLASRG 1803
DB	1447 KTAIVSNNLVKWY--GNFNAVKGVNFHNSKDCFLGVNGAGKTSTFQMLTGENSISGG 1504
QY	1804 EAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVATAGSG 1863
DB	1505 DAYVNGSVKNNWREAGANTGYCQYDAI IKEMSSEETLYMFARIRIGIPEKDIKPKVNAV 1564
QY	1864 LARLGLSWADRPAGTYSGGNKRLKATALALVGPVAVVFLDEPTTGMDPSARRFLWNSLL 1923
DB	1565 IHAIGIGMYASRQIKTYSGGNKRLSLGIAIVGLPVLILLDEPTSGVDPKARRIITWNLN 1624
QY	1924 AVVREGSRVMTLTSMEERCEALCSRLA IWNNGFRCLGSPQHLKGRPAAGHTLTURVPA 1983
DB	1625 RLRLDGTALVUTSHSMECEALCTELAIWVGKFRCYGSCQHIKSRYSGSGYTLRLIKNR 1684
QY	1984 R-SQPAFAFVAAPFPGSLREAHGRLRFQPLPGGRCALARVFGELAVHGAHGVDFSV 2042
DB	1685 NDAEKTKSTIKQTFRGSVKEEHVQLNFDIPRQGD--SWSRLFEKLETVSTSLNWDDYSL 1743
QY	2043 SQTMLSEVFLYFKDQKQKDETEBEQKEAGVGVDPAPGLQH 2082
DB	1744 SQTLEQVIFBFRSDAGVSSDSEFDGASSTGANSRGNQN 1783

Qy 1942 CEALCSRLAIWNGRFRCLGSPQHLGRFAAGHTLTLRVDAAR-SOPAAAFVAEPGSE 2000
 Db 1158 CEALCTELAIWVYGFRCYGCQHIKSRVGSYTLTIRLKNRNDAAKTKSTIKQTFRGSV 1217
 Qy 2001 LREAHGGRLFQLPGGRCALAVFGELAVHGAHGVDEFSVOTMLEEVFLYFSKQDGK 2060
 Db 1218 IKBEHVLQNLNFDIPRDGD-SWSRLEFKLETVSTSLNWDYSLSTQTTLEQVFIERSDAGV 1276
 Qy 2061 DEDTEQKEAGVGVDVPAPGLQH 2082
 Db 1277 SSDSEFDGASSTGSANSRGKQN 1298

RESULT 9
 T15200
 hypothetical protein F12B6.1 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
 C;Accession: T15200
 R;Pauley, A.; Maggi, L.
 submitted to the EMBL Data Library, May 1997
 A;Description: The sequence of C. elegans cosmid F12B6.
 A;Reference number: Z18307
 A;Accession: T15200
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1447 <PAU>
 A;Cross-references: EMBL:AF003138; NID:G2088708; PID:G2088709; PIDN: AAB54153.1; GSPDB:GN
 A;Experimental source: strain Bristol N2; clone F12B6
 C;Genetics:
 A;Gene: CESP:F12B6.1
 A;Map position: 1
 A;Introns: 79/2; 114/3; 177/1; 224/3; 331/1; 345/3; 373/2; 417/2; 464/1; 536/1; 659/2; 6
 C;Superfamily: ATP-binding cassette homology

Query Match 14.7%; Score 1600.5; DB 2; Length 1447;
 Best Local Similarity 27.3%; Pred. No. 6.4e-92;
 Matches 475; Conservative 257; Mismatches 533; Indels 477; Gaps 51;

Qy 446 PLTDLRVVWGSEVYVLO-----DLVERAAVRVLGG-ANPRAGLYLOOMPY 488
 Db 13 PLIDLKYYTGFGLQGSWFSLEQSKSKLSIDRAIMSELNTQTDANLGVTAQOEY 72
 Qy 489 PCYVDDVFLVLSRSLPLFLTLLAWIYVTLTVKAAVREKETRLDTRMAGLSRAVLWL 548
 Db 73 PCTVKDTFNAL--FMPLFLLSIFPSALLVKNVYKEQKIKEQKRAAGLGDVAHFIS 130
 Qy 549 WFTSLCGPFLLSAALLVLVLKGLDILPYSHGVVFLFAAFAVATVTSQSLLSAFTSRAN 608
 Db 131 WGLISLVNLFISVLIIISIKVAKIFDYDTLLFLVLILFFSSIAMSIFFSTLFTNAN 190
 Qy 609 LAAACGGALAYESLYLPVYLVANRDLRPAGRVAASLLSPVAFGFCESALLLEEQGEA 668
 Db 191 IATAATCVLMFVFFIPQLLRTORISPTFNRLS-LLPPTAMGH-CFKLLESFNAMERA 248
 Qy 669 QWNHVG--TRPTADVSLAOSGLLLDLAALYGLIATWYLEAVCPGQIGIEPWNFFRRS 726
 Db 249 TWSDLWENNPVLGI-SVELCMIMLVDTAVFLLAWIISAVAPDGFVGPQLWFFPTLK 307
 Qy 727 YWCGPRPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLE----KRFPGSP-----Q 775
 Db 308 YW-----APGLYKNRVFEVDDEHFDITPNSDSGFSDE 338

Qy 776 P-----NLRLGLSLDFYOGHITAFIHNAGKTKTTLTSLGLFPPSGSASFILGHDVRSRM 830
 Db 339 PTNLJTALDCNLRLYEGQITGLIGHNAGKTKTMSILCGLYAPSSGTAKIYQDRIKTL 398
 Qy 831 AAIPLHGLVCPQYNVLFDMLTVDESHVWFYGRLLKGLSAAVVGPEQDRLLQDVLGVSKOSVQ 890
 Db 399 RRVRDVLGIQFQHNVLFSHLTVSEQLRLFAALKGVPDSELTSQVDEILASVLSKANKL 458
 Qy 891 TRHLSGGMQRKLSVAIAFVGSQVVLDEPTAGVDVPASREGIWEILLKUYEGRTLLSTH 950

Db 1443 -GRFATVAVCLAVDQKCEFCGLLVNGAGKTTTFNILLTQSFASGEAMIGRVDTE--- 1498
Qy 1818 AAHL5MGVCPQSDAIFELLTGREHLELLARLURGVPEAQAQTAGSLARLGLSWYDRPA 1877
Db 1499 --LISIGYCPQFDALMLDLTGRESLEILAAQHGHPENYKA--KAEILILECVGMIAHADKLV 1554
Qy 1878 GTYSGGKRLKATALALVGDPAVVPDPTTGMPSARRFLWNSILLAVVREGS-VMLTS 1936
Db 1555 RFYSGGQKRIKISVGALLAPTQMIIIDPTAGIDPKARREVMELLMCREHNSALMLTS 1614
Qy 1937 HSMDECEALCSRLAIMVNGRCFLGSPQHLKGRFAAGHTLTLRV--PAARSQPAFAFVAA 1994
Db 1615 HSMDECEALCSRIAVNREGSIIAIGSSQELKSLGNYTWTLSLYENQDM-VVQLVQT 1673
Qy 1995 EFPGESELEAHGGR---LRFQLPGGRCALARVFGELAVHGAHGVBDPFSVSTQMLBEVF 2051
Db 1674 RLPNSVLKTTSTNKTNLKWIPEKEDCWSAKFEMVQALAKDLGVKDFILAQSLEETF 1733
Qy 2052 LYFSKDOGKDED 2063
Db 1734 LRLA---GLDED 1742

RESULT 11
S60124
transport protein homolog C48B4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 13-Jan-1996 #sequence_revision 12-Apr-1996 #text_change 16-Aug-2004
C:Accession: S60124; S40724; S40725
R:Kershaw, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: S60124
A:Accession: S60124
A:Molecule type: DNA
A:Residues: 1-1767 <KER>
A:Cross-references: EMBL:Z29117; NID:9439247; PID:g1066912
C:Genetics:
A:Map position: III
A:Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein
F:628-818/Domain: ATP-binding cassette homology <ABCL>
F:645-652/Region: nucleotide-binding motif A (P-loop)
F:764-769/Region: nucleotide-binding motif B
F:1457-1642/Domain: ATP-binding cassette homology <ABC2>
F:1474-1481/Region: nucleotide-binding motif A (P-loop)
F:1586-1591/Region: nucleotide-binding motif B

Query Match 12.9%; Score 1407; DB 2; Length 1767;
Best Local Similarity 25.9%; Pred. No. 1.2e-79;
Matches 485; Conservative 298; Mismatches 665; Indels 424; Gaps 59;

Qy 337 GYSQDAAH--DVGHVGLVTLGRVTECLSLD---KLEAAPSEALVSR-ALQLLAE----- 385
Db 159 GYTSKDAARAKTVDLKKFAERQSAKLKLSVKNSSEEEQLLVRLNDLPLMLNETFCAL 218
Qy 386 HRFWAGVVF-----LG--PESSDPTHEPTDPLGFGHVRKIRMDIDVVT 428
Db 219 NSYAAGVVFDEVDVTNKKLVYRIILGKTPETWHLTETSYNPGFSSGR----- 267
Qy 429 RTNKIRDFDPGPAADPLTLRLVWVGFFVYLDLVERAAVRLSGANPRAGLYLQOMPY 488
Db 268 -----YSRISSPP-----YWTSAFLTFQHAISSEFLSSVQSGAPDLPITRLGPE 313
Qy 489 PCYVDDVFLRLSLPLFLTLAMTYSVTLTVKAVRE--KETRLDTRMAGLSRAVL 545
Db 314 PRYKTSVSAFID-----FFPFIWAFVTFINVIHTREIAENHAVKPYLTAMGLSTFMF 368
Qy 546 WLGFSLCLGPFLLSAALLVLVLKGLDILPYSHGCVGVLFLAAFAVATVQTSLLSAPFS 605
Db 369 YAAHVWMAFLKFFV--IFLCSIIPLTFVMEFVSAAIIVTVMYGLGAVIFGAFAVSPFN 426
Qy 606 RANLAAACGGLAYFSLYLPYVLCVAVNRDL-----PAGGRVAASLLSPV-----ATGFG 654

Db 427 NTNSAIK-----AIIWAGAMIGISYKLRPELDQISSCFLYGLNINGAFALA 473
Qy 655 CESLA-----LLEEQEGAGQWNVGTRPTADVFSLAQVS-GLLLLD-----AALYG 699
Db 474 VEASIDYRRERELNLTNMFNDSLSHFLSGWALVMVIDILMWSIGALVVDHIRTSDAFS 533
Qy 700 LATWY-----LEAVCPGQYGIPEPWNPPFRSYWCGPPPPKSPAPCPTPLDPKV 748
Db 534 LRTLDFEAPEDDENQDTGTAQNTINEQVRNRVRSD-MEMPMMASTSLNPNADSDS 592
Qy 749 LVB-----EAPCGLSPGVSVRSLEKRFPGSPALRGLSLDFYQGHITAFIAGHNGAGKT 802
Db 593 LLEGSTADGARDTARADIIVRNVLKWTSTGERAVDGLSLRAVRGQCSILLGHNGAGKS 652
Qy 803 TTLSILSGLPPPGSGSAFIIIGHDVSRSSMAAIRPHLGVCPOYNVLFDMLTVDHVMFYGRLL 862
Db 653 TTFSSIAGIIRPTNGRITICGYDVGNPEPGETRRHIGMCPQYNPLYDQLTVSEHLKLVVGL 712
Qy 863 KGLSAAVVGPEDRLQDGLVSKQSVQTHLSGQMOKLSVAIAFVGGSQVVLIDEPFTA 922
Db 713 KGAREKDFKQDMKRLSLSDVKLDFKENSKAVNLSGGMKRKLVCVMALIGDSEVLLDEPFTA 772
Qy 923 GVDPASRRGIWELLKYREGRTLLILSTHLDDEALLGDRVAVVAGGRLCCCGSPFLFRRH 982
Db 773 GMDPGARQDYQKLVREKANKRTILLTHYDDEARLGDWFIHSHGKLVASGTYQLKQK 832
Qy 983 LGSYYILTLVKARLPLTTNEKADTMEGSDVTTRQEKNGSQGSRVGTFFQLLALV-----Q 1037
Db 833 FTGYLLTVV-----LDHNGDKRK-----MAVILTDVCT 861
Qy 1038 HWPGARLVEELPHELVLVLPYTCAGHDSFATLPRELD-----TR 1077
Db 862 HYVKEAERGEMHGQIIBIILP--EARKKEFVPLFQALEAIODRNRYRNVNFMNMTLKSQ 919
Qy 1078 LAELRLTGYGISDTSLEEIFLVVEECAADTDMEDSGCQHLCTGIAGLDVTLLKMPQP 1137
Db 920 LATLEMRSGLSLNTLEQVITI-----GDKVKAIASRQ-NSRISHNSR 963
Qy 1138 ETALENGEPAG-----SAPETD-----QSGSPDAGVRVQGHVALTRQQLAALLKRF 1184
Db 964 NASEPSLKAGYDTQSTKSDASQKLMDSQARGPEK-----SGVAKMVAQFISIMRKKFL 1019
Qy 1185 LARRSRGLFAQIVLPALFVGLALVPSLIIVPPGHYPALRLSPMYQAQVSFFSEDPGD 1244
Db 1020 YSRNWAQLFTQVLIPILILGLV-----GSLTTLKSNNT-----D 1054
Qy 1245 PGRARLLEALLQEAAGLEPPVQVSHSRFSAPEVPAEVAKVL-ASG-----NWTPESSPA 1298
Db 1055 QFRS-----LTSGIEPSKVVMWFENGITPEEAAAFKILRKSGGFVLNNTKNP--- 1105
Qy 1299 CQCSQPGARLLPCPAAAGPPPPQAQVTSGEVVQNLTRNLSDFLVTKTYPRLVROGLK 1358
Db 1106 -----LPNITKSLIGEMPATI----- 1122
Qy 1359 TKKWVNEVRVGGPSLGRDPGLPSGQELGRSVBELWALLSPLPGCALDRVLKNTAWAHS 1418
Db 1123 -----GMTN----- 1127
Qy 1419 LDAQDSLKIWFNNKGWH---SMVAFVNRAS-NALTRAHLPPGARHAHSITTLNHLNLT 1474
Db 1128 ---SDNLALFNRYHYVHLPTLISMINEARLTGTVDAEISSGVFLYSKTSN-----SNLL 1180
Qy 1475 KEQLSEALMASSVDVLVSVICVVFAMSFVPASFTVLIERVTRAKHLQMLGGLSPTYLW 1534
Db 1181 PSQI-----IDVLLAPMLILIFAMVTSTFMFLIEERTCQFAHQQLTGISPTIFY 1231
Qy 1535 LGNPLDMCNVLPACIVLIFLAFQORAVVAPANLPALLLLLLLYGWSITPLMYPASFF 1594
Db 1232 SASLIYDGILYSLICLIIFLPMFLAF-----HWMYDHLAIVILFWFLYFSSVPFIYAVSFL 1287
Qy 1595 FSVPTAYVVLTCINLFIINGSMATFVLELFSQKLEQ-VSRILKQVFL-IFPHFCLGR 1652

Db 1288 FQSPSKANVLL--IIQVVISGAALLAVFLIFMIFNIDEWLKSILVNIFMFLPLSPYAFGS 1345
QY 1653 GLIDMVRNQADAFERLQDFQSPRLWEVVGKLLAMVIOGPI-FLLFILL----- 1704
Db 1346 AII-TINTYGMILPSEEL-----MWDHCGKNAWLMGTFGVCSFALFVLQLQFKFYRR 1396
QY 1705 -----IQRHSQLPQPRVRSPLPLGLGDEEDVARERERKVGQATGQDVVLRLNLTQYR 1757
Db 1397 FLSQVMTVRRSHNNVQPMGDLPCV-----ESVSEERVRVHRVNSQNSALVIKDLTKTF- 1451
QY 1758 GORMPAVDRLCLGIPGCGFLGNGAGKSTFPMVGTDTLASGEAVLAGHSVAREPS 1817
Db 1452 -GRFVAVNELCLAVDQKFCFLGNGAGKSTFPMVGTDTLASGEAVLAGHSVAREPS 1507
QY 1818 AAHLMSGVCPOSDAIFELLTGREHLELLARLGRVPEAOVATAGSLARLGLSWADRP 1877
Db 1508 --LISIGVCPQDALMLDLTGRESLEILIAOMHGFENYKA--KAEILIECVGMIAHADKL 1563
QY 1878 GTYSGGNKRKLATALVCDPAVFLDEPTTGMQPSARRFLMNSLLAVVREGS--VMLTS 1936
Db 1564 RFYSGGQKRSVGVALLAPTQMIILDEPTAGIDPKARREVWELLWCKREHSNSALMLTS 1623
QY 1937 HSMERCEALCSRLATVNGRFRCLGSPQHLKGRPAAGHTLTLRV--PAARSOPAAAFVAA 1994
Db 1624 HSMDECEALCSRIAVLNRGSLAIGSSQSLKSLYGNNTMTLSLYEPNQDM--VVQLVQT 1682
QY 1995 EPPGSELREAHGGR---LRFOLPPGRCALARVFGELAVHGAHGVEDFSVSOTMLERVF 2051
Db 1683 RLPNSVLKTTSTNKTNLNKLQIPKEKBCWCSAKFEMVQALAKDLGVKDFILAQSSLETF 1742
QY 2052 LYFSKDOGDKE 2063
Db 1743 LRLA---GLDED 1751

RESULT 12
T42749
A:ATP-binding cassette transport protein homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2004
C:Accession: T42749
R:Wu, Y.C.; Horvitz, H.R.
Cell 93, 951-960, 1998
A:Title: The C. elegans cell corpse engulfment gene ced-7 encodes a protein similar to A
A:Reference number: Z22259; MUID:98297348; PMID:9635425
A:Accession: T42749
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1704 <WUY>
C:Cross-references: UNIPROT:P34358; EMBL:AF049142; NID:g3172340; PIDN:AAC24116.1; PID:g3
A:Note: ced-7
C:superfamily: ATP-binding cassette homology

Query Match 12.9%; Score 1405; DB 2; Length 1704;
Best Local Similarity 25.9%; Pred. No. 1.5e-79;
Matches 486; Conservative 297; Mismatches 668; Indels 422; Gaps 59;

QY 337 GYSWQDAHA--DVHGLVLTGRVTECLSLD---KLEAAPSEAAVSR-ALQLLAE----- 385
Db 92 GYTSKDAAKRTVDLMMKKFAERFQSAKLKLSVKNSSESEQLTLVRLNDLPLMNETFCAI 151
QY 386 HRFWAGVVF-----LG--PEDSSDPTHTPDLGPHGVRIKIRMDIDVVT 428
Db 152 NSYAGGVVFDEVDVNTKKNLIRILGKTPPEETHLITETSYNDYPSGSR----- 200
QY 429 RTNKIRDFWDPGPAADPLTLRLYVVGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPY 488
Db 201 -----YSRIPSSPP-----YWTSAPLTQHAIESFLSSVQSGAPDLPITLRLGLPE 246
QY 489 PCYDDVFLVRLSRSLPLFLTLAWTYSVTLTKAVVRE---KETRLRDTMRAMGLGRAVL 545
Db 247 PRYKTSVSAFID-----FFPPIWAFVTFINVIHITREIAENHAKVPLYTAMGLSTFMF 301

QY 546 WLGFSLCLGPFLLSAALLVLVLKGLDILPYSHPGVFLFLAAFAVATVTSQFLLSARFS 605
Db 302 YAAHVWMAFLKFFV--IFLCSIIPLTFVMEFVSPAALIVTVLMYGLGAVIFGAFVASFPN 359
QY 606 RANLAAACGGIAYLSLYLPVLCVAVDRRL-----PAGGRVAASLLSPV-----AFGRG 654
Db 360 NTNSAIK-----AILVAMGAMIGISYKLRPELDQISSCFYGLNGINGAFALA 406
QY 655 CESLA-----LLEEQGEAGQHNNGTRPTADVFLAQVS-GLLLD-----AALVG 699
Db 407 VEALSDYNNRERELNLTNMFNDSSLHFLSGWALVMIVDILMWSLIGALVVDHIRTSAFDS 466
QY 700 LATWY-----LEAVCPGYGIDPEPWNFPFRRS- YMGCPRPSPKSPAPCPTPLDPK 747
Db 467 LRTLDPFEAPEDDENQTDGVTAAQNTIRINEQVNRVRRSDMEIQMNPMASTSLNPPNADSD 526
QY 748 VLVE-----EAPGLSPGVSRSLERKFPQSPQALRGSLDFYQGHITAPLGHNGAGK 801
Db 527 SLLEGSTEADGARDTARADIIVRNLVKIWTSTGERAVDGLSLRAVRGQCSILLGHNGAGK 586
QY 802 TTTLISLGLFPSPGSAFILGHVDRSSMAAIRPHLGVCPQVNVLFDMLTVDDEHVWFYGR 861
Db 587 STTFSSIAGITIRPTNGRITICGYDVNGEPGETRRHIGMCPQYNPLYDQJTVSEHLKLVG 646
QY 862 LKGLSAAVVGPEQRLLODVGLVSKQSVQTRHLSGGMQRKLSVAIFVGGSQWVILDBPT 921
Db 647 LKGAREDKFQDMKRLSLDVKLDPKENEKANVLSGMRKLCVCNALLIGDSEVILLDEPT 706
QY 922 AGVDPASRRGIWELLKYRREGORTILSTHLLDEAELLGDRVAVAGGRLCCGSLFLR 981
Db 707 AGMDPGARQDVQKLVEREKANRTILLTHYMDEAERLGDWVFIMSHGLVASGTNQYLKQ 766
QY 982 HLGSGYLLTYKARLPLTTNEKADTMEGSDVTROEKKNGSGSRVGPQLLALV----- 1036
Db 767 KFGTYLLTVV-----LDHNGDKRK-----MAVILTDVC 795
QY 1037 QHWVPGARLVEELPHELVLVLPYTGADGSPATFREL-----T 1076
Db 796 THYVKEAERGEMHQIILIP--EARKKEFVPLFOALEAQDRNYSRVNPDMPNLTLS 853
QY 1077 RLAERLTGYIGISDTSLEEIFLKVVEBCAADTDMEDSGCGHLCGTIGAGLDVTLRLKMP 1136
Db 854 QLATLEMRSGLSLNTLFEQVETI-----GDKVDKAIASRQ-NSRSHNS 897
QY 1137 QETALENGEPAG---SAPETD-----QGSGPDVGRVQGWALTRQQLQALLKRF 1183
Db 898 RNASEPSLKPAYDTPQSTKSADSYQKLMDSQARGPEK-----SGVAKMVAQFISIMRKKF 953
QY 1184 LLARRRRGLFAQIVLPALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVSFFSEDAPG 1243
Db 954 LYSRRWAQLFTQVLIPIILGLV-----GSLTTLKSNT-----DQFSVRS----- 995
QY 1244 DPGRRALLEALLQAGLEBPPVQHSHRFSFAPEVPAVAKVL-ASG-----NWTPESSP 1297
Db 996 -----LTPSGIERPSKVVRWFENGITPEEAANPEKILRKSGGFEVLNNTKNP-- 1042
QY 1298 ACQCSQPCARLLPDCPAAACGPPPPQAVTSGEVVQNLTGRLSDFLVKTYFRLVROGL 1357
Db 1043 -----LPNITKSLIGEMPPATI----- 1059
QY 1358 KTKKWNVRYVGGSLGGRDPLPSGBQLGRSVEELWALLSPLPGGALDRVLKNTAWAH 1417
Db 1060 -----GMTN----- 1064
QY 1418 SLDAQDSLKIFWNKNGWH---SMVAFVNRAS-NAILRAHLPPGPARHAHSITTLNHPNL 1473
Db 1065 ---SDNLEALFNMYHYHVLPTLISMINRKLTGTVDAAEISSGVFLYSKSTSN---SNL 1116
QY 1474 TKEQLSEALMASSVDVLVSCVVPFAMSFVASFVTLVLEERTAKHLQMGGLSPTLY 1533
Db 1117 LPSQL-----IDVLLAPMLILIFAMVSTFVMFLIEERTCFQHQOQLTGISPTIF 1167

QY 1971 AAGHTLTILRVP-----AARSQPAA-----1981
 Db 1041 --GNHLELEVPFYNGVKPNEVNSVELENFCQIIQQWLFNVPTQPSLLGLDLEVCIGVSDS 1098
 QY 1990 -----1989
 Db 1099 ITPDTASAEISLSPENVQRIAKPLGNEQRVSTLVPPLPREDVPFDQDLSQQLFRDGGIP 1158
 QY 1990 -----AFVAEPFGPGEISELREAHGRLRFQLPPG--GRCALARVFGELAV 2030
 Db 1159 LPIPAEWLTKKESFALDSPTQSSFFPGATFKSCNGLSIKYLQPLPGEGSLADAFGHLER 1218
 QY 2031 HGAEHGVDEFSVSTQMLEEVPFLYPS 2055
 Db 1219 NRNRLGIAEYSISQSTLTETIFNHPA 1243

RESULT 14
 T46467
 hypothetical protein DKFPz434E1030.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Aug-2004
 C:Accession: T46467
 R:Blum, H.; Baurerachs, S.; Wewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23034
 A:Accession: T46467
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-269 <AAA>
 A:Cross-references: UNIPROT:Q9NT80; EMBL:AL137481
 A:Experimental source: adult testis; clone DKFPz434E1030
 C:Genetics:
 A:Note: DKFPz434E1030.1
 C:Superfamily: ATP-binding cassette homology

Query Match 12.3%; Score 1345; DB 2; Length 269;
 Best Local Similarity 100.0%; Pred.No.5.9e-77;
 Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1674 QFQSPLRWEVVGKLLAMVIOGPLEFTLLQHRSSQLLPQPRVRSLLPLGGEDEDVARE 1733
 Db 1 QFQSPLRWEVVGKLLAMVIOGPLEFTLLQHRSSQLLPQPRVRSLLPLGGEDEDVARE 60
 QY 1734 RERVQCATQGDVLVLRNLTKVYRGQMPAVDRCLGIPGECFGLLVGNAGAKTSTFRM 1793
 Db 61 RERVQCATQGDVLVLRNLTKVYRGQMPAVDRCLGIPGECFGLLVGNAGAKTSTFRM 120
 QY 1794 VTGDTLASRGEAVLAGHSVAREPSAAHLSMGVYCPQSDAIFELLTGREHLELLARLRGVP 1853
 Db 121 VTGDTLASRGEAVLAGHSVAREPSAAHLSMGVYCPQSDAIFELLTGREHLELLARLRGVP 180
 QY 1854 AQVAQTAGSGIARLGLSWYADRAGTYSGGNKRLATALVGDPAVFLDEPTTGMDPS 1913
 Db 181 AQVAQTAGSGIARLGLSWYADRAGTYSGGNKRLATALVGDPAVFLDEPTTGMDPS 240
 QY 1914 ARRFPLNLSLAVVREGRSVMLTSH 1937
 Db 241 ARRFPLNLSLAVVREGRSVMLTSH 264

RESULT 15
 T27121
 hypothetical protein Y53C10A.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004
 C:Accession: T27121
 R:White, S.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z20314
 A:Accession: T27121
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA

Db	923	----	:	----	FIKVNK	929
Qy	1351	R-----LVROGLTKKVVNEVRYGFSLGRDPGLPSQGELGRSVEELWALLSP	PGGA	1404		
Db	930	KGDSKFYKISGGDKNAAILMNIITASAWYL--RDP	SVTK-----	965		
Qy	1405	LDRVLNLTAWAHSLDAQSLKIFWNN-----KGHSMVAFVNRASNAILRAHL	PPGPAP	1459		
Db	966	-----LPHVT-----SRVIMWDPKIKYGLASFFLFEN-----	994			
Qy	1460	HAHSITTLNHPNLNTEQLSEALMASSVDVLVSI	CVFAMSFVPASFTLVLI	EEVRTRA	1519	
Db	995	-----IFFLLVLGAFIQS--TWYLI	EEKICKF	1020		
Qy	1520	KHLQMGGLSPTLYWLGNFILWDMCNVLPACIVVL	FLAFQORAYVAPANL	PALL--LLL	1577	
Db	1021	AHQOYLTGLSTIAYGVWLVWDF-----LLTFFLLTIGIFLS	GVGLQGHIIHV	1072		
Qy	1578	LLYG--WSITPLMYPASFFFSVPSTAVVLT	CINLFINGSMATFVLELFS	DQKLQEV	1635	
Db	1073	IFGGLLFYFAPLVLTISALINTPTRGNFLL--YMFCCIPWLAYS	IVSELHNFPF	IQKYS	1129	
Qy	1636	RIILQOVFLIF-PHFCLGRGLDMVR-----NOAMADAFERLGD	RQFQSPURKEWVVG	KNLLA	1690	
Db	1130	DEIEYGPRIENPSIGFLAGLKMIAALNYPKSGLDK	GHFHLTNLWTEY	GIFFE-----LM	1183	
Qy	1691	MVIOGFLFLFTLLL-----QHRSQLLPQPRVRS	PLLGEED	VARER	1734	
Db	1184	FLFFGGIFL--TILLGCATLKPFRRACFR	TRRSQ--PREKR--YK	GIESCKAVKEE	1236	
Qy	1735	ERVVOGATOD-VLVLNLTNVYRGOMPAVDR	LCIGIPGECFGLLVGNV	GAKTSTFRM	1793	
Db	1237	EQLVQEVDMKETVLVIDGLVKDF--GKFRAVNDLS	ISVGHBEFCFGLM	GANGAKTTFFDI	1294	
Qy	1794	VTGDTLASGEAVLAGSVAREPSAAHLSMGYCP	QSDAIFELLTGREHLELLAR	LGVP	1853	
Db	1295	ITGLTMTPTGGSATIDGHITET-----IHIGYCPQ	PDMALQQISCRQTL	IMAKI	QGYP-1348	
Qy	1854	AQAQTAGGGLARIGLSWYADR	PAGTVSGGNKRLATALALVGDP	AVVFLDETTGM	DPS	1913
Db	1349	NVKEVVFLVDCVGM	SDFYKLVKNGSGQKRSISVGIALMS	RATCIILDEP	TAGIDPR	1407
Qy	1914	ARFLNNSLLAVVREG--SVMLTSHSWECEALCS	RLATMNGRRPCL	SGPHLKG	1968	
Db	1408	ARREIWD-----IHEMEQAKCSIVLTS	HSWECEALCTRIGIRK	GEMTALG	TSQLSKS	1463
Qy	1969	RFAAGHTLTLRVPAARS-QPAAAFVAAE	FFGSELREAHGG--RLRF	PLPPG	RCALARY	2024
Db	1464	QYGNTYNMWTLILNSLEDLESVCVIVSE	MPDVLKTPESSITTSIVWEL	PKSKDK	WSEK	1523
Qy	2025	FGLAVHGAHBGVEDFVSQTMLEEVFLY	PSKQGD	DETEROKEA	2070	
Db	1524	YNOVEVLAKKANAKYMLTOASL	EDTFRILI-----TTEEBEEA	1562		

Search completed: January 3, 2005, 10:55:35
Job time : 66.9746 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 238.389 Seconds
(without alignments)
5068.549 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHCHFFNKPPLPSAGT.....QHPKRVQFLDDPSTAEIVL 2100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10885	99.9	2146	2 Q9BZC4	Q9BZC4 homo sapien
2	10883	99.9	2146	2 Q81ZY2	Q81ZY2 homo sapien
3	10881	99.9	2146	2 Q9NR73	Q9NR73 homo sapien
4	10249	94.1	2008	2 Q9GS58	Q9GS58 homo sapien
5	8431	77.4	2170	2 Q7TNJ2	Q7TNJ2 rattus norv
6	8382.5	76.9	2159	2 Q91V24	Q91V24 mus musculu
7	5629	51.7	2260	2 Q8UVV4	Q8UVV4 gallus gall
8	5603.5	51.4	2261	1 ABC1_MOUSE	P41233 mus musculu
9	5593.5	51.3	2261	1 ABC1_HUMAN	Q95477 homo sapien
10	5510.5	50.6	2201	2 Q8QZB2	Q8QZB2 rattus norv
11	4904	45.0	2269	2 Q8T941	Q8T941 canis fam
12	4904	45.0	2269	2 AAR87835	AAR87835 canis fam
13	4903	45.0	2269	2 Q6T940	Q6T940 canis fam
14	4903	45.0	2269	2 AAR87836	AAR87836 canis fam
15	4902	45.0	2269	2 Q6T942	Q6T942 canis fam
16	4902	45.0	2269	2 AAR87834	AAR87834 canis fam
17	4847.5	44.5	2281	2 O02698	O02698 bos taurus
18	4819.5	44.2	2310	1 ACRH_MOUSE	O35600 mus musculu
19	4819.5	44.2	2310	2 AAR57853	AAR57853 mus muscu
20	4808	44.1	2273	1 ABCR_HUMAN	P78363 homo sapien
21	3551	32.6	2436	2 Q9HC28	Q9HC28 homo sapien
22	3541	32.5	2436	1 ABC2_HUMAN	Q9BZC7 homo sapien
23	3522.5	32.3	2434	1 ABC2_RAT	Q9BZC7 rattus norv
24	3390.5	31.1	2434	1 ABC2_MOUSE	Q9BZC7 rattus norv
25	3351.5	30.8	1771	2 Q76MW7	P41234 mus musculu
26	3351.5	30.8	1771	2 Q76MW7	Q76MW7 homo sapien
27	3279	30.1	1487	2 Q8BPY1	BAA83014 homo sapi
28	2808	25.8	2595	1 AB12_HUMAN	Q8BPY1 mus musculu
29	2715	24.9	1416	2 Q6ZPZ4	Q6ZPZ4 homo sapien
30	2715	24.9	1416	2 BAC98084	BAC98084 mus muscu
31	2714	24.9	2127	2 Q86WI2	Q86WI2 homo sapien

RESULT 1

ID	Q9BZC4	PRELIMINARY;	PRT; 2146 AA.
AC	Q9BZC4;		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)		
DE	ABC transporter member 7.		
GN	Name=ABCA7;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21328888; PubMed=11435699;		
RA	Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,		
RA	Shulenin S., Arnould I., Naudin L., Lafargue C., Rosier M., Jordan B.,		
RA	Mattei M.G., Dean M., Denefle P., Chimini G.;		
RT	"Comparative analysis of the promoter structure and genomic		
RT	organization of the human and mouse ABCA7 gene encoding a novel ABCA		
RT	transporter.";		
RL	Cytogenet. Cell Genet. 92:264-270(2001).		
CC	-1- SIMILARITY: Belongs to the ABC transporter family.		
DR	EMBL; AF328787; AAK00959.1; -		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . . ; IEA.		
DR	GO; GO:0006810; P:nucleotide binding; IEA.		
DR	GO; GO:0006810; P:transport; IEA.		
DR	InterPro; IPR003439; AAA ATPase.		
DR	InterPro; IPR003439; ABC_transporter.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	ProDom; PD000006; ABC_transporter; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER_1; 1.		
DR	PROSITE; PS00893; ABC_TRANSPORTER_2; 2.		
KW	ATP-binding.		
SQ	SEQUENCE 2146 AA; 234305 MW; 2391728D5AD97E75 CRC64;		
Query Match	99.9%;	Score 10885;	DB 2; Length 2146;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 2098;	Conservative 1;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1 PPLEHCHFFNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGEEPRGLSNFNDLSVLSLL 60		
Db	47 PPLEHCHFFNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGEEPRGLSNFNDLSVLSLL 106		
Qy	61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAPQPTKQSPLEPPMDVALLTSLRT 120		
Db	107 ADARTVLGGASAHRTLAGLGLIATLRAARSTAPQPTKQSPLEPPMDVALLTSLRT 166		
Qy	121 ESLGLALGQAQEPHLSLEAAEDLAQELLALRSVELRALQLQRPGTSGPILLEALCS 180		

Q86uq4 homo sapien
Q80t20 mus musculu
Bab86781 homo sapi
Q99758 homo sapien
Q80xt2 mus musculu
Q8mu3 stronglyloce
Q8x420 mus musculu
Q7pzy9 anopheles g
Q9vrg4 drosophila
Q7q7k8 anopheles g
Q6x9g3 mus musculu
Aap73044 mus muscu
Q7g1g5 anopheles g
Q7q5k8 anopheles g

Db	167	ESLGLAQAOEPLHSLLEAAGDLAQELLALRLSLVELRALQORPRGTSGPLELLSEALCS	226
Qy	181	VRGSSSTVPSLNWYEASDLMELVQBPESALPOSSISPAACSELIIGALDLSHPLSRLLWRR	240
Db	227	VRGSSSTVPSLNWYEASDLMELVQBPESALPOSSISPAACSELIIGALDLSHPLSRLLWRR	286
Qy	241	LKPLILGKLFPAPPTPFTRKLMQAVNRTFBEITLLRDVREWEMLGRIPTFFMNDSSNVA	300
Db	287	LKPLILGKLFPAPPTPFTRKLMQAVNRTFBEITLLRDVREWEMLGRIPTFFMNDSSNVA	346
Qy	301	MLQRLQMDGRRQPPGGRDHMEALRSFLDPGSGGYSMQDAHADVGHVLTGRVTEC	360
Db	347	MLQRLQMDGRRQPPGGRDHMEALRSFLDPGSGGYSMQDAHADVGHVLTGRVTEC	406
Qy	361	LSLDKLEAAPSEAAALVSRALQLLAEHRFPWAGVFLGPEDSSDPTTEHPTDLPQGHVRIKI	420
Db	407	LSLDKLEAAPSEAAALVSRALQLLAEHRFPWAGVFLGPEDSSDPTTEHPTDLPQGHVRIKI	466
Qy	421	RMDIDVTRTNKIRDRFWDPCPAADPLTDLRYVWGGFYVLODLVERAAVRVLSGANPRAG	480
Db	467	RMDIDVTRTNKIRDRFWDPCPAADPLTDLRYVWGGFYVLODLVERAAVRVLSGANPRAG	526
Qy	481	LYLOMPYPCYVDDVFLVLSRSLPLFTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL	540
Db	527	LYLOMPYPCYVDDVFLVLSRSLPLFTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL	586
Qy	541	SRAVLWLGWFLSCIGPFLLSAALVLVLKGDILPYSHPGVWFLFLAFAVATVTSQFLL	600
Db	587	SRAVLWLGWFLSCIGPFLLSAALVLVLKGDILPYSHPGVWFLFLAFAVATVTSQFLL	646
Qy	601	SAFSTRANLAACGLAYFSLYLVPLVCVAVWRDLRPAAGRVAASILLSPVAFGFCESIAL	660
Db	647	SAFSTRANLAACGLAYFSLYLVPLVCVAVWRDLRPAAGRVAASILLSPVAFGFCESIAL	706
Qy	661	LEEQEGQWNVGTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPQOYGIPEPWN	720
Db	707	LEEQEGQWNVGTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPQOYGIPEPWN	766
Qy	721	PPFRSRYWCGPRPKSPAPCPTPLDPKVLVEAPPGLSVGVSVRSLEKRPFGSPQPALRG	780
Db	767	PPFRSRYWCGPRPKSPAPCPTPLDPKVLVEAPPGLSVGVSVRSLEKRPFGSPQPALRG	826
Qy	781	LSLDFYOGHITAFIIGHNGAGTKTTLSILSGLFPFGSGSAFTILGHVDRSSMAAIRPHLGVC	840
Db	827	LSLDFYOGHITAFIIGHNGAGTKTTLSILSGLFPFGSGSAFTILGHVDRSSMAAIRPHLGVC	886
Qy	841	POYNVLFDMLTVDHNVFYGRLKGLSAAVVGPEQDRLLQDVLVSKOSVQTRHLSGGMQR	900
Db	887	POYNVLFDMLTVDHNVFYGRLKGLSAAVVGPEQDRLLQDVLVSKOSVQTRHLSGGMQR	946
Qy	901	KLVAIAIPVGGSQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD	960
Db	947	KLVAIAIPVGGSQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD	1006
Qy	961	RVAVVAGRLCCCGSPILFRHLGSGYYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN	1020
Db	1007	RVAVVAGRLCCCGSPILFRHLGSGYYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN	1066
Qy	1021	GSQSRVGTGPOLLALVQHVPVPGARLVEBELPHELVLVLPYTGADHGSFATLFRLEDTRLAE	1080
Db	1067	GSQSRVGTGPOLLALVQHVPVPGARLVEBELPHELVLVLPYTGADHGSFATLFRLEDTRLAE	1126
Qy	1081	LRLTGYGISTLSLEIFLKVVEECAADTMDGSCGQHLCTGTAGLDVTLRLKMPPOETA	1140
Db	1127	LRLTGYGISTLSLEIFLKVVEECAADTMDGSCGQHLCTGTAGLDVTLRLKMPPOETA	1186
Qy	1141	LENCEPAGSAPETDQGGSDVAGVQGWALFRQOLQALLLRLFLARRSRGLFAQVLVP	1200
Db	1187	LENCEPAGSAPETDQGGSDVAGVQGWALFRQOLQALLLRLFLARRSRGLFAQVLVP	1246
Qy	1201	ALFVGLALVFSLLVPPFGHYPALRLSPMTMYGAQVSFFSEADAPGDPGRARLLLEALLQEAGL	1260
Db	1247	ALFVGLALVFSLLVPPFGHYPALRLSPMTMYGAQVSFFSEADAPGDPGRARLLLEALLQEAGL	1306
Qy	1261	EEPPVQHSRRFSAPEVPAEYAKVLASGNWTPSPSPACOCQOPGARRLLPDCPAAAGGP	1320
Db	1307	EEPPVQHSRRFSAPEVPAEYAKVLASGNWTPSPSPACOCQOPGARRLLPDCPAAAGGP	1366
Qy	1321	PPQAVTGSSEVVQNLTGRNLSDFLVKTYPRLVQRLGKTKKWNNEVRYGGFSLGGRDPGL	1380
Db	1367	PPQAVTGSSEVVQNLTGRNLSDFLVKTYPRLVQRLGKTKKWNNEVRYGGFSLGGRDPGL	1426
Qy	1381	PSGOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSLDAQDSLKIWFNNKGNHSMVAF	1440
Db	1427	PSGOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSLSLDAQDSLKIWFNNKGNHSMVAF	1486
Qy	1441	VNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSVVVFAM	1500
Db	1487	VNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSVVVFAM	1546
Qy	1501	SFVPASFTLVLIBERTVTRAKHLQMGGLSPTLVWLGNFMDMNCNLYLPACIVLILFLAQ	1560
Db	1547	SFVPASFTLVLIBERTVTRAKHLQMGGLSPTLVWLGNFMDMNCNLYLPACIVLILFLAQ	1606
Qy	1561	QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSVSPSTAYVVLTCINLFIGINGSMAT	1620
Db	1607	QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSVSPSTAYVVLTCINLFIGINGSMAT	1666
Qy	1621	FVLEFSDQKLEVSRLIKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDQFQSPFLR	1680
Db	1667	FVLEFSDQKLEVSRLIKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDQFQSPFLR	1726
Qy	1681	WEVVGNLLAMVTOGQPLFLFTLLLOHRSQLLOPRVRSPLPLGEEDEDEVARERERVVQG	1740
Db	1727	WEVVGNLLAMVTOGQPLFLFTLLLOHRSQLLOPRVRSPLPLGEEDEDEVARERERVVQG	1786
Qy	1741	ATQGDVLVLRNLTKVYRGQMPAVDRICLGPPECECGLLGVNGAGTKSTFRMVGTGTLA	1800
Db	1787	ATQGDVLVLRNLTKVYRGQMPAVDRICLGPPECECGLLGVNGAGTKSTFRMVGTGTLA	1846
Qy	1801	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGRHLELLARLRGPEAQVAQTA	1860
Db	1847	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGRHLELLARLRGPEAQVAQTA	1906
Qy	1861	GSGLARLGLSWADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDDPSARRFLWN	1920
Db	1907	GSGLARLGLSWADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDDPSARRFLWN	1966
Qy	1921	SLLAUVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV	1980
Db	1967	SLLAUVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV	2026
Qy	1981	PAARSQAAAAFVAAEPFGSELREAHGRLRFPOLPPGRCALARVFGELAVHGAHGVEDF	2040
Db	2027	PAARSQAAAAFVAAEPFGSELREAHGRLRFPOLPPGRCALARVFGELAVHGAHGVEDF	2086
Qy	2041	SVSQTMLLEEVFLYFSKQDQKDETEOKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL	2100
Db	2087	SVSQTMLLEEVFLYFSKQDQKDETEOKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL	2146
RESULT 2			
Q81ZV2	ID	Q81ZV2	PRELIMINARY; PRT; 2146 AA.
AC	Q81ZV2;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	ABC transporter ABCA7.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20549028; PubMed=11095984;		

RA Kaminski W.E., Piehler A., Schmitz G.;
RT "Genomic organization of the human cholesterol-responsive ABC
RT transporter ABCA7: tandem linkage with the minor histocompatibility
antigen HA-1 gene.";

RL Biochem. Biophys. Res. Commun. 278:782-789(2000).

CC -1- SIMILARITY: Belongs to the ABC transporter family.

DR EMBL; AF3111060; AAN04657.1; JOINED.

DR EMBL; AF3111061; AAN04657.1; JOINED.

DR EMBL; AF3111062; AAN04657.1; JOINED.

DR EMBL; AF3111063; AAN04657.1; JOINED.

DR EMBL; AF3111064; AAN04657.1; JOINED.

DR EMBL; AF3111065; AAN04657.1; JOINED.

DR EMBL; AF3111066; AAN04657.1; JOINED.

DR EMBL; AF3111067; AAN04657.1; JOINED.

DR EMBL; AF3111068; AAN04657.1; JOINED.

DR EMBL; AF3111069; AAN04657.1; JOINED.

DR EMBL; AF3111070; AAN04657.1; JOINED.

DR EMBL; AF3111071; AAN04657.1; JOINED.

DR EMBL; AF3111072; AAN04657.1; JOINED.

DR EMBL; AF3111073; AAN04657.1; JOINED.

DR EMBL; AF3111074; AAN04657.1; JOINED.

DR EMBL; AF3111075; AAN04657.1; JOINED.

DR EMBL; AF3111076; AAN04657.1; JOINED.

DR EMBL; AF3111077; AAN04657.1; JOINED.

DR EMBL; AF3111078; AAN04657.1; JOINED.

DR EMBL; AF3111079; AAN04657.1; JOINED.

DR EMBL; AF3111080; AAN04657.1; JOINED.

DR EMBL; AF3111081; AAN04657.1; JOINED.

DR EMBL; AF3111082; AAN04657.1; JOINED.

DR EMBL; AF3111083; AAN04657.1; JOINED.

DR EMBL; AF3111084; AAN04657.1; JOINED.

DR EMBL; AF3111085; AAN04657.1; JOINED.

DR EMBL; AF3111086; AAN04657.1; JOINED.

DR EMBL; AF3111087; AAN04657.1; JOINED.

DR EMBL; AF3111088; AAN04657.1; JOINED.

DR EMBL; AF3111089; AAN04657.1; JOINED.

DR EMBL; AF3111090; AAN04657.1; JOINED.

DR EMBL; AF3111091; AAN04657.1; JOINED.

DR EMBL; AF3111092; AAN04657.1; JOINED.

DR EMBL; AF3111093; AAN04657.1; JOINED.

DR EMBL; AF3111094; AAN04657.1; JOINED.

DR EMBL; AF3111095; AAN04657.1; JOINED.

DR EMBL; AF3111096; AAN04657.1; JOINED.

DR EMBL; AF3111097; AAN04657.1; JOINED.

DR EMBL; AF3111098; AAN04657.1; JOINED.

DR EMBL; AF3111099; AAN04657.1; JOINED.

DR EMBL; AF3111100; AAN04657.1; JOINED.

DR EMBL; AF3111101; AAN04657.1; JOINED.

DR Genew; HGNC:37; ABCA7.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0006810; F:transport; IEA.

DR InterPro; IPR003593; AAA ATPase.

DR Pfam; PF00005; ABC tran; 2.

DR ProDom; PD000006; ABC transporter; 2.

DR SMART; SM00382; AAA; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.

DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.

KW ATP-binding.

SQ SEQUENCE 2146 AA; 234421 MW; 33A128082D7B5BAF CRC64;

Query Match

Best Local Similarity

Matches 2097; Conservative

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLT

Db 47 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLTPEEPGRLSNFNDSLSVRL 106
Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMDVAELLSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMDVAELLSLRT 166
Qy 121 ESLGLALGOAEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSPLLEALCS 180
Db 167 ESLGLALGOAEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSPLLEALCS 226
Qy 181 VRGSPSTVGPISLNWYEASDLMEVQEPESALPDSSLSPACSELICALDHPHLSRLMWR 240
Db 227 VRGSPSTVGPISLNWYEASDLMEVQEPESALPDSSLSPACSELICALDHPHLSRLMWR 286
Qy 241 LKPLILGKLLFAPDTPFTRKLMAQVNRTPFELTLRDVREVMELGPRIFTPMNDSSNVA 300
Db 287 LKPLILGKLLFAPDTPFTRKLMAQVNRTPFELTLRDVREVMELGPRIFTPMNDSSNVA 346
Qy 301 MLQRLLOMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWODAHADVGLHVGTLGRVTEC 360
Db 347 MLQRLLOMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWODAHADVGLHVGTLGRVTEC 406
Qy 361 LSLDKLEAAPSEAAALVSRAQLLAHREHFWAGVVFILGPEDSDPTHEPTDGLPGHVRKI 420
Db 407 LSLDKLEAAPSEAAALVSRAQLLAHREHFWAGVVFILGPEDSDPTHEPTDGLPGHVRKI 466
Qy 421 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 467 RMDIDVTRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526
Qy 481 LYLOQMPYPCVVDVFLRSLPLFLTLAWIYSVTITVKAIVREKETRLRDTWRAMGL 540
Db 527 LYLOQMPYPCVVDVFLRSLPLFLTLAWIYSVTITVKAIVREKETRLRDTWRAMGL 586
Qy 541 SRVILWLGWFLSCILGPFLLSAAALVILKLGDIILPYSHPGVVFLLAFAFATVTSFLL 600
Db 587 SRVILWLGWFLSCILGPFLLSAAALVILKLGDIILPYSHPGVVFLLAFAFATVTSFLL 646
Qy 601 SAFPSTRANLAAACGLAYFSLYLPVYLVCVWRDRLPAGGRVAASLLSPVAFGCGESLAL 660
Db 647 SAFPSTRANLAAACGLAYFSLYLPVYLVCVWRDRLPAGGRVAASLLSPVAFGCGESLAL 706
Qy 661 LEEQEGAGQWENVTGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPNW 720
Db 707 LEEQEGAGQWENVTGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPNW 766
Qy 721 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVEAPPGSLSPGVSVRSLEKRFPGSPQALRG 780
Db 767 PPFRRSYWCGPRPKSPAPCPTPLDPKVLVEAPPGSLSPGVSVRSLEKRFPGSPQALRG 826
Qy 781 LSLDPYQGHITAFILGHNGAGKTTLSILSGLFPSPGGSFATILGHVDRSSMAAIRPHLGYC 840
Db 827 LSLDPYQGHITAFILGHNGAGKTTLSILSGLFPSPGGSFATILGHVDRSSMAAIRPHLGYC 886
Qy 841 POYNVLPDMLTVDEHVMFYGRKGLSAAVCPEDRLLODVLVSKQSVOTRHLGGQW 900
Db 887 POYNVLPDMLTVDEHVMFYGRKGLSAAVCPEDRLLODVLVSKQSVOTRHLGGQW 946
Qy 901 KLSVAIAFVGGSQVVIDEPTAGVDPASRRGIGWELLKYREGRTILSTHHLDEALLGD 960
Db 947 KLSVAIAFVGGSQVVIDEPTAGVDPASRRGIGWELLKYREGRTILSTHHLDEALLGD 1006
Qy 961 RVAVVAGRLCCCGSPFLRRHLSGGYYLTLVKARLPLTTNEKADTMEGSVDTROEKN 1020
Db 1007 RVAVVAGRLCCCGSPFLRRHLSGGYYLTLVKARLPLTTNEKADTMEGSVDTROEKN 1066
Qy 1021 GSQSRVCTPOLLAIVQHWVFGARLVEELPHELVLVLPYTCAGHDSFATLRELDTRAL 1080
Db 1067 GSQSRVCTPOLLAIVQHWVFGARLVEELPHELVLVLPYTCAGHDSFATLRELDTRAL 1126
Qy 1081 LRLTYGIGSDTSLEEIPLKVVVECAADTDMEDGSGQHLCGTGAGLDVTLRLKMPQETA 1140

Qy 1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLTPEEPGRLSNFNDSLSVRL 60

Db 1127 LRLTGYGSDTSLEEIFLKVVVECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1186
Qy 1141 LENCEPAGSAPETDQSGPDAGVQGVQWALTRQIQALLLKRFLLARRSRGLFAQIVLP 1200
Db 1187 LENCEPAGSAPETDQSGPDAGVQGVQWALTRQIQALLLKRFLLARRSRGLFAQIVLP 1246
Qy 1201 ALFVGLALVFLSVLPFGHPALRLSPTMYGAQVSFFSEDAPGDPGRARLLEALLQAGL 1260
Db 1247 ALFVGLALVFLSVLPFGHPALRLSPTMYGAQVSFFSEDAPGDPGRARLLEALLQAGL 1306
Qy 1261 BEPPVQHSSHRFSAPAEVPAEYAKVLASGNWTPESPSPACQSQOPGARRLLPDCPAAAGGP 1320
Db 1307 BEPPVQHSSHRFSAPAEVPAEYAKVLASGNWTPESPSPACQSQOPGARRLLPDCPAAAGGP 1366
Qy 1321 PPPQAVTGSSEVQNLTRGNLSDFLVKTYPRLVQGLKTKKWNEVRYGGFSLGGRPGL 1380
Db 1367 PPPQAVTGSSEVQNLTRGNLSDFLVKTYPRLVQGLKTKKWNEVRYGGFSLGGRPGL 1426
Qy 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSDAQDSLKTFWNNKGWHSWAF 1440
Db 1427 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSDAQDSLKTFWNNKGWHSWAF 1486
Qy 1441 VNRSNAILRAHLPFGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSVVVFAM 1500
Db 1487 VNRSNAILRAHLPFGPARHAHSITTLNHPNLTKQELFEAALMASSVDVLVSVVVFAM 1546
Qy 1501 SFVPAFTVLVIERVTRAKHLQMGGLSPTLYWGNFLMDMCMNYLPACIVLVLIFFAQ 1560
Db 1547 SFVPAFTVLVIERVTRAKHLQMGGLSPTLYWGNFLMDMCMNYLPACIVLVLIFFAQ 1606
Qy 1561 QRAVAPANLPALLLLLYGWSITPLMYPASFPSPSTAYVVTCTINLFIGINGSMAT 1620
Db 1607 QRAVAPANLPALLLLLYGWSITPLMYPASFPSPSTAYVVTCTINLFIGINGSMAT 1666
Qy 1621 FVLELFDQKLQEVSRILKQVFLFPFHCLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1680
Db 1667 FVLELFDQKLQEVSRILKQVFLFPFHCLGRGLIDMVRNQAMADAFERLGDROFQSPFLR 1726
Qy 1681 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQPRVRSPLPGEEDVARERERVQ 1740
Db 1727 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPQPRVRSPLPGEEDVARERERVQ 1786
Qy 1741 ATQGDVLVRLNTKVYRGORMPADVRLCLGIPPGCEGLLGVNCGAGKTSFRWVTGDTLA 1800
Db 1787 ATQGDVLVRLNTKVYRGORMPADVRLCLGIPPGCEGLLGVNCGAGKTSFRWVTGDTLA 1846
Qy 1801 SRGEAVLAGHSVAREPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1860
Db 1847 SRGEAVLAGHSVAREPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA 1906
Qy 1861 GSGLARLGLSWADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLWN 1920
Db 1907 GSGLARLGLSWADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPTTGMDPSARRFLWN 1966
Qy 1921 SLLAVVREGSVMLTSSMECEALCSRLATMNGRFRCLGSPHLLKGRFAAGHTLRLV 1980
Db 1967 SLLAVVREGSVMLTSSMECEALCSRLATMNGRFRCLGSPHLLKGRFAAGHTLRLV 2026
Qy 1981 PAARSQPAAPAAVFAEFPGSELREAHGRLRFLQPLPGGRCALARYFGEALVGAHGEHVEDF 2040
Db 2027 PAARSQPAAPAAVFAEFPGSELREAHGRLRFLQPLPGGRCALARYFGEALVGAHGEHVEDF 2086
Qy 2041 SVSOTMLEEVFLYFSKQKDEDETEOKEAGVGDPAFGLOHPKRVSQFLDDPSTAETVL 2100
Db 2087 SVSOTMLEEVFLYFSKQKDEDETEOKEAGVGDPAFGLOHPKRVSQFLDDPSTAETVL 2146

RESULT 3

Q9NR73

PRELIMINARY; PRT; 2146 AA.

ID Q9NR73

AC Q9NR73;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Macrophage ABC transporter.
GN Name=ABCA7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20334305; PubMed=10873640;
RA Kaminski W.E., Orso E., Diederich W., Klucken J., Drobnik W.,
RA Schmitz G.;
RT "Identification of a novel human sterol-sensitive ATP-binding cassette
RT transporter (ABCA7)";
RL Biochem. Biophys. Res. Commun. 273:532-538(2000).
CC -/- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF285794.1; -.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005524; F:ATP binding; TAS.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; TAS.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
KW ATP-binding.
SQ SEQUENCE 2146 AA; 234468 MW; 679B16EBD275FF0D CRC64;

Query Match 99.9%; Score 10881; DB 2; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2098; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PPLEHECHPPNKLPSAGTVPWLQGLICNVNNTCPQLTPGEBPGLSNFNSLSRLL 60
Db 47 PPLEHECHPPNKLPSAGTVPWLQGLICNVNNTCPQLTPGEBPGLSNFNSLSRLL 106
Qy 61 ADARTVLGGASARTLAGLKLATLRAARSTAQPTKOSPPLPMLDVALLTSLLRT 120
Db 107 ADARTVLGGASARTLAGLKLATLRAARSTAQPTKOSPPLPMLDVALLTSLLRT 166
Qy 121 ESLGLALGQAEPLHSLLEAAEDLAQELLALRSIVELRALLQRPRTSGPLELSEALCS 180
Db 167 ESLGLALGQAEPLHSLLEAAEDLAQELLALRSIVELRALLQRPRTSGPLELSEALCS 226
Qy 181 VRGFSSTVGPLSNWYEASDLMELVQEPESALPDSSLSACSELIGALDSHPLSRLWRR 240
Db 227 VRGFSSTVGPLSNWYEASDLMELVQEPESALPDSSLSACSELIGALDSHPLSRLWRR 286
Qy 241 LKPLILGLKLLFAPDTPPTRKLMAQVNTFELTLRLRDVREVWMLGPRIETFMNDSSNVA 300
Db 287 LKPLILGLKLLFAPDTPPTRKLMAQVNTFELTLRLRDVREVWMLGPRIETFMNDSSNVA 346
Qy 301 MLQRLQMDQEGRRQPRGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
Db 347 MLQRLQMDQEGRRQPRGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406
Qy 361 LSLDKLEAAPSEAAALVSRAQLLAHRFWAGVVFVLPGEDESDPTHEPTDPLGPGHVRKI 420
Db 407 LSLDKLEAAPSEAAALVSRAQLLAHRFWAGVVFVLPGEDESDPTHEPTDPLGPGHVRKI 466
Qy 421 RMDIDVVRTNKLDRDFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 480
Db 467 RMDIDVVRTNKLDRDFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 526
Qy 481 LYLQOMPYPYCVDDVFLVRLSRSLPLFLTLAWTYSVTLTKAVVREKETRLRDTMRMAGL 540
Db 527 LYLQOMPYPYCVDDVFLVRLSRSLPLFLTLAWTYSVTLTKAVVREKETRLRDTMRMAGL 586
Qy 541 SRAVLWLGWFLSCIGPFLLSAALLVLVILKGLDILPYSHPGVWFLFLAFAVATVTSQFLL 600
Db 587 SRAVLWLGWFLSCIGPFLLSAALLVLVILKGLDILPYSHPGVWFLFLAFAVATVTSQFLL 646

QY 601 SAFFSRANLAAACGGAYFSLYPYVLCVAMDRLPAGGRVAASLLSPVAFPGCESLAL 660
DB 647 SAFFSRANLAAACGGAYFSLYPYVLCVAMDRLPAGGRVAASLLSPVAFPGCESLAL 706
QY 661 LEEQEGAGHNVGTRPTADVPSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPNW 720
DB 707 LEEQEGAGHNVGTRPTADVPSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPNW 766
QY 721 PFRFRSYWCGPPRPPKSPACPTPLDPKVLVEBAPGLSPGVSRSLKEKFPSPALRG 780
DB 767 PFRFRSYWCGPPRPPKSPACPTPLDPKVLVEBAPGLSPGVSRSLKEKFPSPALRG 826
QY 781 LSLDPYQGHITAFGLHNGAGKTTTILSGLLPPSGGSFAFILGHVDVRSMAAIRPHLGYC 840
DB 827 LSLDPYQGHITAFGLHNGAGKTTTILSGLLPPSGGSFAFILGHVDVRSMAAIRPHLGYC 886
QY 841 POYNYVLFDMLTVDHWFVGRLLKGLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSGMOR 900
DB 887 POYNYVLFDMLTVDHWFVGRLLKGLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSGMOR 946
QY 901 KLSVALAFVGGSOVVLDPBTPAGVDPASRRGIWELLKYREGRTILSTHLLDEAELLGD 960
DB 947 KLSVALAFVGGSOVVLDPBTPAGVDPASRRGIWELLKYREGRTILSTHLLDEAELLGD 1006
QY 961 RVAVVAGRLCCCGSPLFLRRHLGSGYLLTVKARLPLTTNEKADTMEGSDVTTRQEKGN 1020
DB 1007 RVAVVAGRLCCCGSPLFLRRHLGSGYLLTVKARLPLTTNEKADTMEGSDVTTRQEKGN 1066
QY 1021 GSGGRVGTPLLALVQHVVPGARLVEELPHELVLVLPYTGADHSGSFATLPRELDTRLA 1080
DB 1067 GSGGRVGTPLLALVQHVVPGARLVEELPHELVLVLPYTGADHSGSFATLPRELDTRLA 1126
QY 1081 LRLTGVSIDTSLLEELFLKVEECADTDMEGSCGHLCTGIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTGVSIDTSLLEELFLKVEECADTDMEGSCGHLCTGIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAGSAPETDQSGSDVAVGRVQGWALTQQQLALLKRFLLARRSRGLFAQIVLP 1200
DB 1187 LENGEPAGSAPETDQSGSDVAVGRVQGWALTQQQLALLKRFLLARRSRGLFAQIVLP 1246
QY 1201 ALFVGLALVFSILVPPFGHYPALRLSPNTMYGAQVFPSEDAPDGPGRARLLLEALLQ 1260
DB 1247 ALFVGLALVFSILVPPFGHYPALRLSPNTMYGAQVFPSEDAPDGPGRARLLLEALLQ 1306
QY 1261 BEPPVQHSRHSAPAEVPAEVAKVLASGNWTPESPACQSPGARRLLPCPAAAGGP 1320
DB 1307 BEPPVQHSRHSAPAEVPAEVAKVLASGNWTPESPACQSPGARRLLPCPAAAGGP 1366
QY 1321 PPQAVTSGEVVQNLTRNLSDFLVKTYPRLVROGLTKKWVNEVRYGGFSLGGRDPGL 1380
DB 1367 PPQAVTSGEVVQNLTRNLSDFLVKTYPRLVROGLTKKWVNEVRYGGFSLGGRDPGL 1426
QY 1381 PSGQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSLEADQSLKIWFNNKGWHSWAF 1440
DB 1427 PSGQELGRSVEELWALLSPLPGCALDRVLKNTAWAHSLEADQSLKIWFNNKGWHSWAF 1486
QY 1441 VNRSNAILRAHLPGRPARHAHSITTLNHPNLNTEQISEAALMASSVDVLVSIQVWAM 1500
DB 1487 VNRSNAILRAHLPGRPARHAHSITTLNHPNLNTEQISEAALMASSVDVLVSIQVWAM 1546
QY 1501 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPVACIVLLIFLAFO 1560
DB 1547 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLNFLWDMCNVLPVACIVLLIFLAFO 1606
QY 1561 QRAYVAPANLPALELLLYGWSITPLMPYASFFSVSTAYVTLTINLFIINGSMAT 1620
DB 1607 QRAYVAPANLPALELLLYGWSITPLMPYASFFSVSTAYVTLTINLFIINGSMAT 1666
QY 1621 FVLELFSDOKLOEVRILKQVFLIPPHFLGRGLIDMVRNQAMADAFERLGRQFQSPILR 1680
DB 1667 FVLELFSDOKLOEVRILKQVFLIPPHFLGRGLIDMVRNQAMADAFERLGRQFQSPILR 1726

QY 1681 WEVVGKNLLAMVIOGPLFLFTLLQHRSQLPQPRVSLPLIGDEEDVARERERVQ 1740
DB 1727 WEVVGKNLLAMVIOGPLFLFTLLQHRSQLPQPRVSLPLIGDEEDVARERERVQ 1786
QY 1741 ATQGDVLVLRNLTKVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLVLRNLTKVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTTSTFRMVTGDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDALFELLTGREHLELLARLGVPEAQVAQTA 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDALFELLTGREHLELLARLGVPEAQVAQTA 1906
QY 1861 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVVFLEDBTTTGMDFSARRFLWN 1920
DB 1907 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVVFLEDBTTTGMDFSARRFLWN 1966
QY 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTILRV 1980
DB 1967 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTILRV 2026
QY 1981 PAARSOPAAAFAVAEPFPGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAEGHVEDF 2040
DB 2027 PAARSOPAAAFAVAEPFPGSELREAHGGRRLRFOLPPGRCALARVFGELAVHGAEGHVEDF 2086
QY 2041 SVSQTMLBEVFLYFSDQKQEDTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2100
DB 2087 SVSQTMLBEVFLYFSDQKQEDTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146
RESULT 4
Q96S58 PRELIMINARY; PRT; 2008 AA.
AC Q96S58;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ABCA-SSN.
GN Name=ABCA7/ABCA-SSN;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21255283; PubMed=11355874;
RA Tanaka A., Ikeda Y., Abe-Dohmae S., Arakawa R., Sadanami K.,
RA Kigera A., Nakagawa S., Nagase T., Aoki R., Kioka N., Amachi T.,
RA Yokoyama S., Ueda K.;
RT "Human ABCA1 Contains a Large Amino-Terminal Extracellular Domain
Homologous to an Epitope of Sjogren's Syndrome";
RL Biochem. Biophys. Res. Commun. 283:1019-1025(2001)
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB055390; BAB62294.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:nucleotide binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2008 AA; 218616 MW; 226FF85C24230B90 CRC64;
Query Match 94.1%; Score 10249; DB 2; Length 2008;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1978; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	121	ESLGALGQAQEP	PLHSLLEAAEDLAQELLALRSVLREALLOR	PRGTSGLPLELSEALCS	180
Db	29	ESLGALGQAQEP	PLHSLLEAAEDLAQELLALRSVLREALLOR	PRGTSGLPLELSEALCS	88
Qy	181	VRGFSSTVPSLNWYEASDL	MELVQBPESALPOSSLPACSELIGALDSHPLSRLLWRR	240	
Db	89	VRGFSSTVPSLNWYEASDL	MELVQBPESALPOSSLPACSELIGALDSHPLSRLLWRR	148	
Qy	241	LKPLILGKLFPADPTFTR	KLMAQVNRTEFELTLRLDRVREWMLGPRIFTFMNDSSNVA	300	
Db	149	LKPLILGKLFPADPTFTR	KLMAQVNRTEFELTLRLDRVREWMLGPRIFTFMNDSSNVA	208	
Qy	301	MLQRLQWQDEGRQPR	PGGRDHWEALRSFLDPGSGGYSWQDAHADVGHVLTIGRVTETC	360	
Db	209	MLQRLQWQDEGRQPR	PGGRDHWEALRSFLDPGSGGYSWQDAHADVGHVLTIGRVTETC	268	
Qy	361	LSLDKLEAAPSEAAALV	SRALQLLAEHRFPWAGVFLGPEDSDPTHEPTDLPQGHVRIKI	420	
Db	269	LSLDKLEAAPSEAAALV	SRALQLLAEHRFPWAGVFLGPEDSDPTHEPTDLPQGHVRIKI	328	
Qy	421	RMDIDVTRTNKIRDR	FWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	480	
Db	329	RMDIDVTRTNKIRDR	FWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	388	
Qy	481	LYLOOMPYPYVDDVFL	VRLSRSLPLFLTLAWIYSVTILTVKAVREKETRLDRTRAMGL	540	
Db	389	LYLOOMPYPYVDDVFL	VRLSRSLPLFLTLAWIYSVTILTVKAVREKETRLDRTRAMGL	448	
Qy	541	SRAVLWLGWFLSCLG	PFLLSAALVLVLKLGDIILPYSHPGVVFLLAFAVATVTSQFLL	600	
Db	449	SRAVLWLGWFLSCLG	PFLLSAALVLVLKLGDIILPYSHPGVVFLLAFAVATVTSQFLL	508	
Qy	601	SAFFSRANLAAACGL	LAYFSLPYLVCVWRDLRDPAGGRVAAASLLSPVAFPGCESLAL	660	
Db	509	SAFFSRANLAAACGL	LAYFSLPYLVCVWRDLRDPAGGRVAAASLLSPVAFPGCESLAL	568	
Qy	661	LEEGEGQWQHNVT	TRPTADVFLSAQVSGLLILLDAALYGLATWYLEAVCPQOYGIPEPWN	720	
Db	569	LEEGEGQWQHNVT	TRPTADVFLSAQVSGLLILLDAALYGLATWYLEAVCPQOYGIPEPWN	628	
Qy	721	PPFRSRYWCGPRPK	SPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRRPGSPQALRG	780	
Db	629	PPFRSRYWCGPRPK	SPAPCTPLDPKVLVEEAPPGLSPGVSVRSLEKRRPGSPQALRG	688	
Qy	781	LSLDFYQGHITAF	LGHNGAGKTTTLSILSGLFPSPGSAFILGHDRVSSMAAIRPHLGVC	840	
Db	689	LSLDFYQGHITAF	LGHNGAGKTTTLSILSGLFPSPGSAFILGHDRVSSMAAIRPHLGVC	748	
Qy	841	PQYNVLFDMLTVD	DEHWFYGRKGLSAAVQPEQDRILLQDVLVSKOSVQTRHLSGGMQR	900	
Db	749	PQYNVLFDMLTVD	DEHWFYGRKGLSAAVQPEQDRILLQDVLVSKOSVQTRHLSGGMQR	808	
Qy	901	KLVAIAFVGSQV	ILDEPTAGVDPASRRGIGWELLKLYREGRTLILSTHLDDEAELLGD	960	
Db	809	KLVAIAFVGSQV	ILDEPTAGVDPASRRGIGWELLKLYREGRTLILSTHLDDEAELLGD	868	
Qy	961	RVAVAGGRCLCCGS	PLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN	1020	
Db	869	RVAVAGGRCLCCGS	PLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN	928	
Qy	1021	GSQSRVGTTPOLLAL	VQHWVPGARLVBELPHELVLVLPYTGADHGSFATLIFRELDTRLAE	1080	
Db	929	GSQSRVGTTPOLLAL	VQHWVPGARLVBELPHELVLVLPYTGADHGSFATLIFRELDTRLAE	988	
Qy	1081	LRLTGYSIDTSL	LEEIFLKVVEEACADTDMEDGSCGQHLCTGIAGLDVTLLRKMPPOETA	1140	
Db	989	LRLTGYSIDTSL	LEEIFLKVVEEACADTDMEDGSCGQHLCTGIAGLDVTLLRKMPPOETA	1048	
Qy	1141	LENPEAGSAPET	DQSGDPDVGVRVQGWALTRQQLQALLKRLFLARRSRRLGFLAQIVLP	1200	
Db	1049	LENPEAGSAPET	DQSGDPDVGVRVQGWALTRQQLQALLKRLFLARRSRRLGFLAQIVLP	1108	

Qy	1201	ALFVGLALVLSL	IVPPFGHYPALRLSPMTYGAOVSPFSEDAPGDPGRARLLLEALLQEAGL	1260			
Db	1109	ALFVGLALVLSL	IVPPFGHYPALRLSPMTYGAOVSPFSEDAPGDPGRARLLLEALLQEAGL	1168			
Qy	1261	EEPPVQHSSHRFSA	PEVPAEVAKYLASGNMTSPSPACQSQPGARRLLPDCPAAAGGP	1320			
Db	1169	EEPPVQHSSHRFSA	PEVPAEVAKYLASGNMTSPSPACQSQPGARRLLPDCPAAAGGP	1228			
Qy	1321	PPQAVTGSSEVQNT	LGRNLSDFLVKTYPRLVROGLKTKKWVNEVRYGGFSLGGRDPGL	1380			
Db	1229	PPQAVTGSSEVQNT	LGRNLSDFLVKTYPRLVROGLKTKKWVNEVRYGGFSLGGRDPGL	1288			
Qy	1381	PSGQELGRSVEEL	WALLSLPLPGGALDRVLKNLTAWAHSLEDAOSLKTWFNNKGWHSWVAF	1440			
Db	1289	PSGQELGRSVEEL	WALLSLPLPGGALDRVLKNLTAWAHSLEDAOSLKTWFNNKGWHSWVAF	1348			
Qy	1441	VNRASNAILRAH	LPFGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSI	1500			
Db	1349	VNRASNAILRAH	LPFGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSI	1408			
Qy	1501	SFVPASFTLVLI	IBERVTRAKHLQMLGGLSPTLWLNFLWDMCNLYVPACIVVLIFLAQ	1560			
Db	1409	SFVPASFTLVLI	IBERVTRAKHLQMLGGLSPTLWLNFLWDMCNLYVPACIVVLIFLAQ	1468			
Qy	1561	QRAYVAPANL	PALLLLLLYGWSITPLMYPASPFESVPSTAYVVLTCINLFIGINGSMAT	1620			
Db	1469	QRAYVAPANL	PALLLLLLYGWSITPLMYPASPFESVPSTAYVVLTCINLFIGINGSMAT	1528			
Qy	1621	FVLELSDQKLE	QVSRILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDROFQSPLR	1680			
Db	1529	FVLELSDQKLE	QVSRILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDROFQSPLR	1588			
Qy	1681	WEVVGKLLAMW	IQGPLFLFTLLQHRSQLLPQPRVRSPLLGEDEEDVAREHERVVOG	1740			
Db	1589	WEVVGKLLAMW	IQGPLFLFTLLQHRSQLLPQPRVRSPLLGEDEEDVAREHERVVOG	1648			
Qy	1741	ATQGDVLVRLN	LTKVYRGQRMPAVDRCLGIPGCECFGLLVNGAGKTSFRMVTGDTLA	1800			
Db	1649	ATQGDVLVRLN	LTKVYRGQRMPAVDRCLGIPGCECFGLLVNGAGKTSFRMVTGDTLA	1708			
Qy	1801	SRGEAVLAGHS	VAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA	1860			
Db	1709	SRGEAVLAGHS	VAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTA	1768			
Qy	1861	GSGLARLGLSW	ADRPAGTSGGNKRKLATALALVGPDPVFLDEPTTGMDPSARRFLWN	1920			
Db	1769	GSGLARLGLSW	ADRPAGTSGGNKRKLATALALVGPDPVFLDEPTTGMDPSARRFLWN	1828			
Qy	1921	SLLA	VREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLV	1980			
Db	1829	SLLA	VREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLRLV	1888			
Qy	1981	PAARSQAAAAF	VAAEFPFSGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHGEVDF	2040			
Db	1889	PAARSQAAAAF	VAAEFPFSGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHGEVDF	1948			
Qy	2041	SVSQTMLEEV	FLYFSKQKQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL	2100			
Db	1949	SVSQTMLEEV	FLYFSKQKQKDEDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL	2008			
RESULT 5							
Q7TNJ2	ID	Q7TNJ2	PRELIMINARY;	PRT; 2170 AA.			
AC	Q7TNJ2;						
DT	01-OCT-2003	(TrEMBLrel. 25, Created)					
DT	01-OCT-2003	(TrEMBLrel. 25, Last sequence update)					
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)					
DE	ATP-binding cassette transporter sub-family A member 7.						
GN	Name=ABCA7;						
OS	Rattus norvegicus (Rat).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.						

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wister; TISSUE=Placental;
RA Sasaki M., Nada S., Yamauchi A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB097814; BAC81426.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti...; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2170 AA; 237718 MW; 003C8DF70B8744CE CRC64;

Query Match 77.4%; Score 8431; DB 2; Length 2170;
Best Local Similarity 76.5%; Pred. No. 0;
Matches 1646; Conservative 154; Mismatches 272; Indels 80; Gaps 12;

QY 1 PLEHHECHFPNKPSPAGTVPWLOGLICNVNNTCPQLTGCPEEGRSLNFNDLSVRL 60
DB PLEHHECHFPNKPSPAGTVPWLOGLICNVNNTCPQLTGCPEEGRSLNFNDLSVRL 60
QY 61 ADARTVLGASAHRTIAGLGLIATLRAARSTAQP-----QPTKQSPLEPMLDVAELT 116
DB ADAHTVLGSHSTQDMLAALGLKIPVLRVSGSNWQESNQPAKQ-----SVTELEK 159
QY 117 LRTSLGLALGOAEPHLSLEAEADLAQELLALRSVLVELRALLQRPRTSGPSELISE 176
DB ILQRASLETVLGAQDSMEKFSDATRTVAQELLTLPSELVELRALLRRPRGSAGSELISE 219
QY 177 ALCSVRGPGSTVGPSLNWYEASDLMLVCOEPESALPDSSLPACSELIGALDSHPLSRL 236
DB ALCSKTVGSPGLSLNWTAEANQINFMGPPELAPTLPDSSLPACSEFVGLDDHPVSRLL 279
QY 237 LWRLKPLILGLKLPADPTPTFKLMAQVNRITFEELTLIRDVREVMGLGPRIFTFMNDS 296
DB LWRLKPLILGLKLPADPTPTFKLMAQVNRITFEELTLIRDLHELWGLVGPQIFNMNDS 339
QY 297 SNVAMLQRLQMDGRRQPRGGRDHMEALRSFLDPGSGGYSWQADHADVGHVLTGR 356
DB TNVAMLQKLLDVEGTGWQOQTPKQKLEAIRDFLDPSRGRYNNQEAHADMGRLAEILQG 399
QY 357 VTECLSLDKLEAAPSEALVSALOLLAHRRFWAGVVLGPPEDSDPTHEHTP-DLGPQH 415
DB ILECVSLDKLEAVPSEALVSALBELGERRWAGIVFLSPBHPDSDSEPPSTPTTGPQH 459
QY 416 VRIKIRMDIDVTRNKIRDRFWDGPAADPLTLRYVWGGFVYLQDLVRAAVRLSGA 475
DB LVKIRMDIDVTRNKIRDFKWDGPSADPLMDLRYVWGGFVYLQDLLEQNAVRLSGR 519
QY 476 NPRAGLYLQMPYPCYVDVDFLRLVLSRLPLFLTLAWIYSVTLTKAVVREKETRLRDTM 535
DB DSRAGLYLQMPHPCYVDVDFLRLVLSRLPLFLTLAWIYSVALTKAVVREKETRLRDTM 579
QY 536 RAMGLSRVAVLWGLFSLCGLPPLSAALLVVLKLDILPYSHPGVWFLFLAAFAVATV 595
DB RAMGLSRVAVLWGLFSLCGLPPLSAALLVVLKLDILPYSHPGVWFLFLAAFAVATVA 639
QY 596 QSFLLSAFFSRANLAAAACGLAYFSLPYLVLCVAVRDLRAGGRVAAASLLSPVAFGFC 655
DB QSFLLSAFFSRANLAAAACGLAYFALYLPYLVLCVAVRDLRPLUGGLAVLSLSPVAFGFC 699
QY 656 ESLALLEEGGGAQWHNYGTRPTADVFSLAQVSGLLLDLAALYGLATWYLEBAVCPQYGI 715
DB LAMVAQGPFLFTLLQLHRRNRLPQPKSRLLPPLPGEDEEDVVRERERTKGTGQDVLV 1819

Db 700 ESLALLEEGGGAQWHNLGTGPAEDVFSLAQVSAFLLLDAVTYGLALWYLEAVCPQYGI 759
QY 716 PEPWNPFRRSYWCGRPPPKSPAPCPTPLDPKVLVEEAPGLSPGVSVLSLEKRPFGSQ 775
Db 760 PEPWNPFRRSYWCGRPPPKSVLAPAPQDPKVLVEEPPGLVPGVSVIRGLKHPFGSQ 819
QY 776 PALRGLSLDFYQGHITAFILGHNGAGKTTTSLISLGLFPPSGGSAFTLGHDRSSMAAIRP 835
Db 820 PALRGLNLDFYEGHITAFILGHNGAGKTTTSLISLGLFPPSGGSAFTLGHDRSSMAAIRP 879
QY 836 HLGVCPCYNVLPDMLTVDEHVMFYGRLLKGLSAAVVPGEQRLLDQVLGVSKSVQTRHLS 895
Db 880 HLGICPCYNVLPDMLTVDEHVMFYGRLLKGLSAAVVPGEQRLLDQVLGVSKSVQTRHLS 939
QY 896 GGMQRKLSVAIAFVGSQVWILDEPTAGVDPASRRGIWELLKLYREGRTLLIETHLDEA 955
Db 940 GGMQRKLSVAIAFVGSQVWILDEPTAGVDPASRRGIWELLKLYREGRTLLIETHLDEA 999
QY 956 ELLGDRVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNE-KADTDMEGSVDT 1014
Db 1000 ELLGDRVAVVAGSLCCCGSPFLRRHLGSGYYLTLVKSSQSLVTHDLKGDTE-----DP 1054
QY 1015 ROEKXNGSG-----SRVGTPO-----LLALVQHVWVPGARLVEEL 1049
Db 1055 RREKXSGSEGTADTVLTRDGHRSQVPAADPAVPVTPSAALILELVQHVWVPGARLVEEL 1114
QY 1050 PHELVLPYTCAGHDGSPATLRELDRLAELRTGYGISTLSLEEIFLKVVVEECANDTD 1109
Db 1115 PHELVLPYAGALDGSFATVQELDQQLERLGLTYGISTNLBEIFLKVVVEEAAHAGE 1174
QY 1110 MEDSCGQHLCTGTAGLDVTLRLKMPQORTALENGE-----PAGSAPETDQSGP 1159
Db 1175 GGDPRQOHLT-----ATPQHTGPEASVLENGELAKLVLDPOAPKGSATTTAQ----- 1224
QY 1160 DAVGRVQWALTROQLQALLKRFLLARRRRGLFAQIVLPALFVGLALVFLIIVPPFGH 1219
Db 1225 -----VOGWTLTCCQRLALLKRFLLARRRRGLFAQIVLPALFVGLALVFLIIVPPFGH 1279
QY 1220 YPALRLSPMYGCAQVFESEDPAGDPGRARLEALLQEALEEPVQVSHSR----- 1271
Db 1280 YPQLSLPAMYGPQVFFSEDPADPNRMKLEALLGBAGLQDPSPVQSGKSGSECTHSL 1339
QY 1272 ---FSAPVPAEVAEVAKVLASGNWTPESPSPACQSPGARRLLPDCPAAAGPPPPQAVTG 1328
Db 1340 ACYFTVPEVPDVASILASGNWTPDPSPPACQSPGARRLLPDCPAGAGPPPPQAVTG 1399
QY 1329 SGVVQNLTRNLSDFLVKTYPRLVRQGLTKKWNEVRYGFSGLGRDPDPLPSGREVVR 1388
Db 1400 FGEVQNLTRNVSDFLVKTYPRLVRQGLTKKWNEVRYGFSGLGRDPDPLPSGREVVR 1459
QY 1389 SVEELWALLSPGGLADRVLKNLTAWAHSILDAQDSLKIWFNNKWHSMVAFVNASNAI 1448
Db 1460 TVAENRALLSPQGNLTDRILNLTQWALGDARNLSKIWFNNKWHSMVAFVNASNGL 1519
QY 1449 LRAHLPPGPAPAHASITTLNHPNLTKQLSEALMASSVDVLSICVVFAMSFVPASPT 1508
Db 1520 LRAFLPSSVRAHASITTLNHPNLTKQLSEALMASSVDVLSICVVFAMSFVPASPT 1579
QY 1509 LVLIEERVTRAKHQLMGGLSPTLYWLGFLWDMCNLYLPACIVVLIIFLAFOQRAYVAPA 1568
Db 1580 LVLIEERVTRAKHQLVSGLPQTLWLGFLWDMCNLYLVAICIVVLIIFLAFOQRAYVAPA 1639
QY 1569 NLPALLLILLYGWSITPLMYPASFFSPVSTAYVLTCLNLTGINSMATFVLELSD 1628
Db 1640 NLPALLLILLYGWSITPLMYPASFFSPVSTAYVLTCLNLTGINSMATFVLELSD 1699
QY 1629 QKLQEVSRILKQVFLIPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPLRWEVGNKL 1688
Db 1700 QNLQEVSRILKQVFLIPHFCLGRGLIDMVRNQAMADAFERLGRQFQSPLRWEVGNKL 1759
QY 1689 LAMVQGPFLFTLLQLHRRSOLLPQPRVRSLLPLGEEDEEDVVRERERTKGTGQDVLV 1748
Db 1760 LAMVAQGPFLFTLLQLHRRNRLPQPKSRLLPPLPGEDEEDVVRERERTKGTGQDVLV 1819

Qy 1749 LRNLTKVYRGORPAVDRLCLGIPPGCFGLGUNGAGKSTSTFRMTVGTDLASRGEAVLA 1808
 Db 1820 LRDLTKVYRGORSAPVDHLCIGIPPGCFGLGUNGAGKSTSTFRMTVGTDLPSGGEAVLA 1879
 Qy 1809 GHSVAREPSAAHLSMGYCPQSDAIFELLCTGREHLEILARLGRVPEAOVAQTAGSLARLG 1868
 Db 1880 GHVVAQPSAAHRSMGYCPQSDAIFDLITGREHLEILFARLGRVPEAOVAQTALSGVRLG 1939
 Qy 1869 LSWYADRPAGTYSGNKRKLATALLALVGDPAVFLDEPTTGMDSARRFLWNSLLAVRE 1928
 Db 1940 LPSYADRPAGTYSGNKRKLATALLALVGDPAVFLDEPTTGMDSARRFLWNSLLAVRE 1999
 Qy 1929 GRSVWLTSHSMECEALCSRLA IMVNGRFRCLSGPQHLKGRPAAGHTLTLRVPAARSQPA 1988
 Db 2000 GRSVWLTSHSMECEALCSRLA IMVNGRFRCLSGPQHLKGRPAAGHTLTLRVPAARSQPA 2059
 Qy 1989 AAFVAAEPGSELREAHGRLRFOLPPGRCALARVFEGLAVHGAEGHVEDFSVSQTMLE 2048
 Db 2060 IAFIVTTFPAELREVHSRURFOLPPGGCTLARVFEGLAAQKAGHVEDFSVSQTMLE 2119
 Qy 2049 EVFLYFSKDGKDRDTEQKEAGYGVDPAPFGLQHPKRVSOFLDDPSTAETVL 2100
 Db 2120 EVFLYFSKDGKEEGSQETETREVS-TPGLQHPKRVSRFLEDPSSVEITVI 2170

RESULT 6
 Q91V24
 ID Q91V24 PRELIMINARY; PRT; 2159 AA.
 AC Q91V24;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ATP-binding cassette transporter sub-family A member 7.
 GN Name=Abca7;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=O1129, and DBA/2;
 RX MEDLINE=21328888; PubMed=11435699;
 RA Broccardo C., Osorio J., Luciani M.-F., Schriml L.M., Prades C.,
 RA Shulenin S., Arnold J., Naudin L., Lafargue C., Rosier M., Jordan B.,
 RA Mattei M.G., Dean M., Denefle P., Chimini G.;
 RT "Comparative analysis of the promoter structure and genomic
 RT organization of the human and mouse Abca7 gene encoding a novel ABCA
 RT transporter";
 RL Cytogenet. Cell Genet. 92:264-270(2001).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 CC EMBL; AF287142; AAK56863.1; -;
 DR EMBL; AF287141; AAK56862.1; -;
 DR HSSP; P58301; 1F2U.
 DR MGD; MGI:1351646; Abca7.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:000524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
 DR GO; GO:000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; ABC_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA_2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00435; PROXIDASE_1; UNKNOWN_1.
 KW ATP-binding.
 SQ SEQUENCE 2159 AA; 236881 MW; CD2BE3FE0D8B822B CRC64;

Query Match 76.9%; Score 8382.5; DB 2; Length 2159;

Best Local Similarity 76.4%; Pred. No. 0;
 Matches 1635; Conservative 160; Mismatches 277; Indels 69; Gaps 11;
 Qy 1 PPLEHCHFPNPKPLPSAGTVPWLQGLICNVNNTCPQLTPGEPGRLSNFNDLSVRL 60
 Db 47 PPLEHCHFPNPKPLPSAGTVPWLQGLICNVNNTCPQLTPGEPGRLSNFNDLSVRL 106
 Qy 61 ADARTVLGASAHRTLAGLKGKLIATRAASTAQP---OptKOSPLEPPMLDVAELLTS 116
 Db 107 ADTRTVLGGHSIQMDLDALGKLPVLRVGGGARPOESDQPTSQG-----SVTKLEK 159
 Qy 117 LRTESIGLALGAQOAPFLHSLLEAAEDLAQELALRSLVELRALLORPRTGSPLELSE 176
 Db 160 ILQASILDVPLGAQASMRKFSDAIRDLAQELTLPSLMELRALLRRPGSAGSLELVSE 219
 Qy 177 ALCSVRGSPSTVGPSLNWYEASDLMELVGPEPSALPDSSLSPACSELIGALDSHPLSRL 236
 Db 220 ALCSVTKGPSGGLSLNWYEANQLNEFMGPPEVAPALPDNSLSPACSEFVGTLDHPVSR 279
 Qy 237 LWRRLKPLILGKLIAPDTPFTKMAOVNRTTEELTLARDVREVMELGPRTFTMNDS 296
 Db 280 LWRRLKPLILGKLIAPDTPFTKMAOVNRTTEELTLARDVREVMELGPRTFTMNDS 339
 Qy 297 SNVAMLQRLIQMDQEGRRQPRPGCRDHMEALRSFLDPGSGYSGWQDAHADVHGLVGTGR 356
 Db 340 TNVAMLQRLLDVGGTGTQROQTTPRAOKKLEAIKDFLDPGRGYSWREAHADMGLAGILG 399
 Qy 357 VTECLSLDKLEAPSAALVSRLAQLLAHRFWAGVWVFLGPEDSSDPTHEPTDPLGRHV 416
 Db 400 MMECVSLDKLEAVPSBEALVSRLALELGGRRILWAGIVFLSPHEPLDPSELSPALSPGHL 459
 Qy 417 RIKIRMDIDVVTNKIRDFWDPGPAADPLTLRYVMGSGFVYLQDLVEAAVRLSGAN 476
 Db 460 RFIKIRMDIDVVTNKIRDFWDPGPAADPLTLRYVMGSGFVYLQDLVEAAVRLSGAN 519
 Qy 477 PRAGLVIQMPYPCYVDDDFLRLVLSRLPLFLTLAMIYSVTLTKAVREKETRLRDTMR 536
 Db 520 SRTGLVIQMPHPCYVDDDFLRLVLSRLPLFLTLAMIYSVTLTKAVREKETRLRDTMR 579
 Qy 537 AMGLSRVLMWFLSCGLGPFLLSAALLVLVLKGLDILPYSHPGVVFLLAAFAVATVQ 596
 Db 580 AMGLSRVLMWFLSCGLGPFLLSAALLVLVLKGLNLPYSHPVVIFLFAAFATAVQ 639
 Qy 597 SFLLSAFFSSANLAAACGGGLAVESLYLPPVLCVARDRLPAGRVAASLLSPVAFGCE 656
 Db 640 SFLLSAFFSSANLAAACGGGLAVESLYLPPVLCVARDRLPAGRVAASLLSPVAFGCE 699
 Qy 657 SLALLBEQSGEAGQHNVTPTADVFSLAQVSGLLILDALYGLATWYLEAVCPGQYGP 716
 Db 700 SLALLBEQSGEAGQHNVTPTADVFSLAQVSGLLILDALYGLATWYLEAVCPGQYGP 759
 Qy 717 EPWNFPFRSRYWCGPRPPKSPAPCPPTLPDKVLVESAPGLSPGVSVRSLEKPPGPOP 776
 Db 760 EPWNFPFRSRYWCGPRPPKSPAPCPPTLPDKVLVESAPGLSPGVSVRSLEKPPGPOP 819
 Qy 777 ALRGLSLDFYQGHITAFLGHNGAKTTTILSGLEPPSGSASFLIGHDVSSMAATRP 836
 Db 820 ALQGLNLDFYEGHITAFLGHNGAKTTTILSGLEPPSGSASFLIGHDVSSMAATRP 879
 Qy 837 LGVCPQNVNLFDMLTVDHVMVYGRKLGSLAAVVGPEQDRLLDQVGLSVKSVQTRHLSG 896
 Db 880 LGICPQNVNLFDMLTVDHVMVYGRKLGSLAAVVGPEQDRLLDQVGLSVKSVQTRHLSG 939
 Qy 897 GMQRKLSVAIAFVGGSOVILDEFTAGVDPASRRGIWELLKLYREGRTLLSTHHLDEAE 956
 Db 940 GMQRKLSVAIAFVGGSOVILDEFTAGVDPASRRGIWELLKLYREGRTLLSTHHLDEAE 999
 Qy 957 ILGDRVAVAGGRLCCCGSPLELRHLLGSCYVLTAVKARLPLATNE-KATDMEGSDTR 1015
 Db 1000 ILGDRVAVAGGRLCCCGSPLELRHLLGSCYVLTAVKARLPLATNE-KATDMEGSDTR 1054
 Qy 1016 QEKNGSQSGS-----RVGTPQLLALVQHWVPGARLVBELP 1050

Db 1055 REKSDGNGRTSDTATFTRGTSKSNQAPAGVPIPTPSTARILELVQHVPGQALVEDLP 1114
QY 1051 HELVLVLPYTGADHGSFATLRELDTRLAELRLTYGISTDTSLIEIFLKVVECAADTDM 1110
Db 1115 HELVLVLPYAGALDGSFAMVFELOQLELGLTYGISTDTNLEIFLKVVED--AHREG 1172
QY 1111 EPGSCQHCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGSPDAVGRVQGWAL 1170
Db 1173 GDSRPQLHRT-----CTPQPTGPEASVLENGELA-----PQLAPNA-AQVQGWTL 1219
QY 1171 TFOQLQALLKFLARSRGLFAQIVLPALFVGLALVFSLIVPPFGHPALRLSPMY 1230
Db 1220 TCQURALLKFLARSRGLFAQIVLPALFVGLALVFSLIVPPFGHPALRLSPMY 1279
QY 1231 GAQVFFSDAPGDCGRALLLEALLQEALEBPVQVHSHR-----FSAPEVPA 1279
Db 1280 GPVQFFSDEADGDPNRMKLEALLGEAGLQPSQDKDARSECTHSLACVFTVPEVPP 1339
QY 1280 EVAKVLASGNWTPSPSPACQSQCGARLLPDCPAAAGPPPPQAVTSGSEVWQNLTR 1339
Db 1340 DVASILASGNWTPSPSPACQSQCGARLLPDCPAGAGPPPPQAVAGLGEVWQNLTR 1399
QY 1340 NLSDFLVKTYPLVRCQGLKTKWNEVRVYGGFSLGGRDPLSGQBLGSRVBEELWALLSP 1399
Db 1400 NVSDFLVKTYPLVRRGLTKKWVDEVRYGGFSLGGRDPLFTGHEVVRTIAEIRALLSP 1459
QY 1400 LFGGALDRVLKNTAWHSLDQDSLKIWNKNGHSMVAFVNRASNAILRAHLPPGPAP 1459
Db 1460 QFGNALDRILNLTQWGLDARNLSKIWNKNGHSMVAFVNRANGLLHALLSPGVR 1519
QY 1460 HAHSITTLNHPNLTKQELSEALMASSVDVLVSVICVVPFSPASFTLVLIBERVTRA 1519
Db 1520 HAHSITTLNHPNLTKQELSEALMASSVDVLVSVICVVPFSPASFTLVLIBERVTRA 1579
QY 1520 KHLQMGSLPTLYMGLNFMDCNLYVPACIVLIFLAFQRAYVAPANLPALELILL 1579
Db 1580 KHLQVSGLPOTLYMGLNFMDCNLYVAVCTVWFIFLAFQRAYVAPENLPALELILL 1639
QY 1580 YGWSITPLMYPASFPSPSTAYVVLTCINLFIGINGSMATVLELFDQKQVSRILK 1639
Db 1640 YGWSITPLMYPASFPSPSTAYVVLTCINLFIGINGSMATVLELFDQKQVSRILK 1699
QY 1640 QVFLIPPHFCGLRGILDMVRNQAMADAFERLGDROFQSPLRWEVVGKLLAMVIGQPLF 1699
Db 1700 QVFLIPPHFCGLRGILDMVRNQAMADAFERLGDROFQSPLRWDIIGKLLAMVIGQPLF 1759
QY 1700 LFTLLQHRSQLLPQVRVSLPLLEGEDEVDARERVVQATQGDVLVRLNLTQVYRGQ 1759
Db 1760 LFTLLQHRNRLLPQVRVSLPLLEGEDEVDARERVVQATQGDVLVRLNLTQVYRGQ 1819
QY 1760 RMPAVDRLCGLIPGCEFCGLLVGNGAGKTSTFRMTGDTLASRGAVLGHSVAREPSAA 1819
Db 1820 RMPAVDRLCGLIPGCEFCGLLVGNGAGKTSTFRMTGDTLASRGAVLGHSVAREPSAA 1879
QY 1820 HLSMGYCPQSDAIFELLTGREHLELRLRGVPEAQVATAGTAGSLARGLSWYADRPAGT 1879
Db 1880 HRSMGYCPQSDAIFELLTGREHLELRLRGVPEAQVATAGTAGSLARGLSWYADRPAGT 1939
QY 1880 YSGGNKRLATALVGPVFLDEPTTMDPSARRFLWNSLLAVVREGSRVMTLSHM 1939
Db 1940 YSGGNKRLATALVGPVFLDEPTTMDPSARRFLWNSLLAVVREGSRVMTLSHM 1999
QY 1940 BECEALCSRLA LTVNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQAPAAFAVEFPGS 1999
Db 2000 BECEALCSRLA LTVNGRFRCLGSPHLLKGRFAAGHTLTLRVPAARSQAPAAFAVEFPGS 2059
QY 2000 ELREAAGGRRLPQLPPGRCALARYFGLA VHAEGHGVDEFVSQTMLEEVFLYPSKQDQ 2059
Db 2060 ELREAAGGRRLPQLPPGRCALARYFGLA VHAEGHGVDEFVSQTMLEEVFLYPSKQDQ 2119
QY 2060 KDEDTEBQKAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2100
Db 2120 BEEESSRQ--EABEEBVSXPGROHPKRVSRFLEDPPSSVETMI 2159

RESULT 7
Q8UVV4
ID Q8UVV4 PRELIMINARY; PRT; 2260 AA.
AC Q8UVV4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ATP-binding cassette transporter 1.
GN Name=ABCA1;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RA Attie A.D., Brooks-Wilson A.R., Walker D., McManus B.,
RA Gray-Kellar M.P., MacDonald M.L.E., Roomp K., Tebon A., Zhang L.-H.,
RA Mulligan J., Sensen C., Bitgood J.J., Cook M.E., Kastelein J.J.P.,
RA Hayden M.R.,
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF362377; AAL56247.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2260 AA; 254070 MW; 19D137F342F98662 CRC64;

Query Match 51.7%; Score 5629; DB 2; Length 2260;
Best Local Similarity 49.8%; Pred. No. 0;
Matches 1117; Conservative 363; Mismatches 588; Indels 174; Gaps 25;
QY 1 PPLEHHECHFPNKPLPSAGTVPWLGGLICVNNNTCFPQLTPCEEPGRLSNFDLSVRL 60
Db 47 PPYQHQECHFPNKAMPAGSAGTLFWIQTICNANNPCFRTPTGESPIGVNFASVRL 106
QY 61 ADARTVLGGASAHRTLAGLGLIATLRAA----- 89
Db 107 SDAKRLLYSQDTSIKDVQKVLAKRLKGNSSGLDLKLRDPLVDNETFSDPLRHVSNP 166
QY 90 -----RSTAQFPQTKQSPLEPP 106
Db 167 SSAAVELLDAEVLNQLKIVSGYRIQLRDLNCSALSEFLTIONRSVANDSEAFCLTPKE 226
QY 107 MLDVAEL-----LTSLLATESLGLALGOAQLHSLLEAAEDLAQELALRLVLR 158
Db 227 TLHAEALAFRANLPLKPLQREIFNNSLRDLSETVEALRDSLGKLVKELLSWSMDR 286
QY 159 ---ALLORPRGTSGPLEL---LSEALCSVRGSPSTVGPSSLNWEASDLMLVVG----QEP 208
Db 287 QEVMLTNNVNASSTQIYQAVSRIVCGHPGGGLKIKSLNWNEDNNYKALFGGNSTEDD 346
QY 209 ESALPDSSLSPACSELIGALDHPRLSLRLWRRLKPLILGKLLFADPTFTKRLMAQVNR 268
Db 347 VTNFYDNTTTPYCNELMKNLESSPLSIIRALKPLLGKLVLYTPTPAIRKINAEVNR 406
QY 269 FEETLLRDVREVMELGPRIFTFNDSSNVAMLQRLIQ-----MQDEGR 313
Db 407 FOELGVFRLDGLGMMWEISPKIWTFMESQEMDLIRTLKSKALMDLHLHPASNWTVEDVAR 466
QY 314 ---RQPRPGRDHMEALRSLFDPGSGGYSDQADAHVGLVGLTGRVTECLSLDKLEAAP 370

[2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RX Qiu Y., Cavelier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies
revealing novel regulatory sequences.";
RL Genomics 73:66-76(2001).
RN [3]
RN INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not
ABCG1 via an LXR-independent pathway.";
RL J. Lipid Res. 43:952-959(2002).
RN [4]
RN DOWN-REGULATION BY ENDOTOXIN.
RX PubMed=1277468; DOI=10.1194/jlr.M300100-JLR200;
RA Khovidhunkit W., Moser A.H., Shigenaga J.K., Grunfeld C.,
RT "Endotoxin down-regulates ABCG5 and ABCG8 in mouse liver and ABCA1 and
ABCG1 in J774 murine macrophages: differential role of LXR.";
RL J. Lipid Res. 44:1728-1736(2003).
CC -!- FUNCTION: CAMP-dependent and sulfonylurea-sensitive anion
transporter. Key gatekeeper influencing intracellular cholesterol
transport (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed in adult tissues. Highest
levels are found in pregnant uterus and uterus.
CC -!- INDUCTION: Down-regulated by endotoxins (LPS) or cytokines (TNF
and IL-1) in J774 macrophages. The down-regulation by endotoxin in
macrophages is not likely to be mediated by the liver X
receptor/retinoic X receptor (LXR/RXR).
CC -!- DOMAIN: Multifunctional polypeptide with two homologous halves,
each containing an hydrophobic membrane-anchoring domain and an
ATP binding cassette (ABC) domain.
CC -!- PTM: Phosphorylation on Ser-2054 regulates phospholipid efflux (By
similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X75926; CA53530.1; ALT_INIT.
DR EMBL; AF287263; AAG39073.1; ALT_INIT.
DR MGD; MGI:99607; Abcal.
DR GO; GO:0008203; P:cholesterol metabolism; IDA.
DR GO; GO:0030301; P:cholesterol transport; IDA.
DR GO; GO:0042158; P:lipoprotein biosynthesis; IMP.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Phosphorylation; Transmembrane; Transport.
FT TRANSMEM 26 42 Potential.
FT TRANSMEM 640 656 Potential.
FT TRANSMEM 690 706 Potential.
FT TRANSMEM 717 733 Potential.
FT TRANSMEM 749 765 Potential.
FT TRANSMEM 771 787 Potential.
FT TRANSMEM 1041 1057 Potential.
FT TRANSMEM 1351 1367 Potential.
FT TRANSMEM 1661 1677 Potential.
FT TRANSMEM 1708 1724 Potential.
FT TRANSMEM 1737 1753 Potential.
FT TRANSMEM 1775 1791 Potential.
FT TRANSMEM 1854 1870 Potential.
FT NP_BIND 933 940 ATP (Potential).

FT	NP_BIND	1946	1953	ATP (Potential).
FT	MOD_RES	1042	1042	Phosphoserine (by PKA) (By similarity).
FT	MOD_RES	2054	2054	Phosphoserine (by PKA) (By similarity).
FT	CARBOHYD	14	14	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	98	98	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	151	151	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	161	161	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	196	196	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	244	244	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	292	292	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	337	337	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	349	349	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	400	400	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	478	478	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	489	489	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	521	521	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	820	820	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1144	1144	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1294	1294	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1453	1453	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1499	1499	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1504	1504	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	1637	1637	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	2044	2044	N-linked (GlcNAc) (Potential).
FT	CARBOHYD	2238	2238	N-linked (GlcNAc) (Potential).
FT	CONFLICT	1567	1568	Missing (in Ref. 2).
FT	CONFLICT	2024	2024	Missing (in Ref. 2).
SQ	SEQUENCE	2261	AA, 253972	MM, F0C3C5F1CFE34F1 CRC64;

Query Match 51.4%; Score 5603.5; DB 1; Length 2261;
Best Local Similarity 49.6%; Pred. No. 0;
Matches 1118; Conservative 363; Mismatches 577; Indels 195; Gaps 27;

QY	1	PPLHHEHCHFNKPLPSAGTVPWLOGLICNVNNTCFPQLTGCEEPGRISNFNDSLSRL	60
DB	47	PPYEQHCHFNKAMPAGTLPWVGQIICNANNPCFRYPTEGAPGVGNFKSVSR	106
QY	61	ADARTVLGGASNAHRTLAGLGLKIATRAAR	96
DB	107	SDAQELLYSQDTSIKDMHVKVLMRLQIKPNLSNKLQDFLVNDNETSGFLQHLNLSLR	166
QY	97	PTKOSPLBPPM	112
DB	167	STVDSLLQANVGLQKVFQYQLVGLHSLCNGSKLEBIIQLGDAEVSALCGLPRKLLDAE	226
QY	113	LL	155
DB	227	RVLRYNMDILKPVVTKLNSLSTHLPQHLA	279
QY	156	ELR---ALLQPRGTSGPLEL---LSEALCSVRGSPSTVGPSLNWYASDLMLVGV---	205
DB	280	DMRQEVMTLVNNSSTQIYQAVSRIVCGHPEGGLKIKSLNWNWYNNYKALFGGNT	339
QY	206	QEPESALPDSLSLSPACSELI GALSHPRLRLRLKPLIIGKLFPADPTPTTRKLMQAV	265
DB	340	EEDVDVTYDNTTTPVCNDLMKNLESSPSLR IIKWALKPLLVGKILYTPDTPATRVQMAV	399
QY	266	NRTPEELTLDDRVWEMGLPRTFTFNDSSNVAMLQRLLOMQ--DGRQRPGRGRDH	323
DB	400	NKTQELAVFHDLGEMWELSPQIWTFMENSQEMDLVRLTLLDSRNDQFWEQKLGDLWT	459
QY	324	MEALRSFL-----DRSGGYSWQDAHADVGHVLTGLRVTECLSDKLEAAPEAL	375
DB	460	AQDIWAFIAKNPVEDVQSPNGSVYTWREAFNETNQAIQIISRFMECVNUNKLEPIETVRL	519
QY	376	VSRALQLAEHRFWAGVVFLGPEBSSDPTHTPDLG--PGHVRIRKIMDDVVTRTKI	433
DB	520	INKSMELDERKFWAGIVFTG-----ITPDSVELPHHVKKYKIRMDINVERTNKI	569
QY	434	RDRFWDPGPADPLTDLRYWGGFVYVYLVQDIYVERAAVRVLSGANPRAGLYLQMPYPCYVD	493
DB	570	KDGYMDPGPRADPFEDMRVYWGGAFLQDVVEQAIIRVLTSEKKTGVYVQMPYPCYVD	629

Qy	494	DVFLRVLSRSLPLPLTLTAWIYSVTLTWKAVVREKETELRDMRAMGLSVAVLWGLWFLSC	553
Db	630	DIFLVRNVRSGMPLFTMTLAWIYSVAVITKSIIVYEKEARLKETMTIMGLDNGILLFWSFWVSS	689
Qy	554	LGPFLLSAALLVVLKGLDILPYSHPGVVFLLFAAPAVATVTSQFLLSAFPSFRANLAAAC	613
Db	690	LIPULVSGALLVTLKLGNNLFPYSDPSVVFVLSVFANVTILOCFLLTSTLFSFRANLAAAC	749
Qy	614	GGLAYFLSYLPYVLVCVAVWRDLRPLAGGVAASLLSPVAFGFCBSLALLEBQEGGAQWHNV	673
Db	750	GGIIFYTLXPYVLVCVAQDYGVPFSIKIPASLLSPVAFGCEFYALFESQIGVGQWDNL	809
Qy	674	GTRPT-ADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGOYGIPENWNPFRPRSYWCGPR	732
Db	810	FESVPEDEGPNLTTAVSNMFLPDTFLYGVMTWYIEAVFPQYGIPIRPWYFPTCKTSYMFGE	869
Qy	733	PPKSPACPPLD-PKVLVEEAPGLSPGVSVRSLEKRFPGSQOPALRGHLSLOPYOQHIT	791
Db	870	IDEKSHPGSQKGVSEICMBEETHPLRLGVSIQNLVKYRDGMKVAVDGGLALNPFYEGQIT	929
Qy	792	AFLGHNGAGKTTTILSGLFPFPGSGSAFILGHNVRSMAAIRPHLGVCQYNNVLFOMLT	851
Db	930	SFLGHNGAGKTTTMSLTGLFPPTSGTAYILGKDIRSEMSSIRQNLGVCQHNVLFDMLT	989
Qy	852	VDEHVWIFYRGLKLSAAVWGPEDORLLODVL-VSKOSVQTRHLSGMQORKLSVAIAFVG	910
Db	990	VEEHIFVVARLKLSEKHVAEMEQMALDVGLPPLSKSKTSQLSGGMQOKLSVALAFVG	1049
Qy	911	GSQVVIDDEPTAGVDPASRRGIWELLILKYREGRTLILSTHHLDEAELLGDRVAVVAGRL	970
Db	1050	GSKVVIDDEPTAGVDYPSRRGIWELLILKYRQGRITILSTHMHDEADILGRIAIISHGL	1109
Qy	971	CCGSGPLFLRRHLGSGVYTLVLVKARLPLTTN-----EKAOTDMEGSVDTROEK	1018
Db	1110	CCVGSSFLPKNLQGTGYTLTLVKDVESSLSCHNSSSTVSLCKEDSVSQSSDAGLGS	1169
Qy	1019	KNGQSRVGTQPQLLALVQHVVPGCARLVEELPHELVLVLPYTGADHGSFATLPRELDTRL	1078
Db	1170	DHESDTLTI DVSAISNLIRKHVS EARLVEDIGHELTVVL PYEAAKEGAFVELPHEIDRL	1229
Qy	1079	ABLRLTYGSDTSLBIFLKVVECAADTMDG-----SCGHLCTGI	1123
Db	1230	SDLGISSYGISSETTLBIFLKVAEESGVDAFTSDGTLPARNNRRAFGDKOSC-LHPFTD	1288
Qy	1124	AGLDVTLRLKMPPOETALENGEPAGSAPETDOGGPDVAG--RVQGWALTRQQLQALLK	1181
Db	1289	DAVD-----PNDSDID-----PESRETDLLSGMDKGSYQLKGWLTQQQFVALLWK	1335
Qy	1182	RFLIARRSRGLFAQIIVLPALFVGLALVFSILVPPFGHYPALRLSPMTYGAQVSPFSEDA	1241
Db	1336	RLLIARRSRKGFPAQIIVLPVAVFCIALVFSILVPPFGKYPSELPQPMWYNEQYTFVSND	1395
Qy	1242	PGDGRARLLEALLQERAG-----LEBPPVQHSHSRPSPAEVPAEVAKVLASGWT	1291
Db	1396	PEDMGTOELLNALTDPGFGTRCMEGNPIDPTPCLAGEEDWTISPVFPQSIVDLFPQNGWT	1455
Qy	1292	PESPSPACQSPCARLRLDPCAAAGPPPPQAVTSGSEVVQNLTCRNLSDFLVKTYPR	1351
Db	1456	MKNPSPACQCSDDIKIKNWLVCPVPGAGLPPPPQKQKTADILQLTLGRNLSDYLKVTYQ	1515
Qy	1352	LVRQGLKTKWVNRVRYGGFSLG-GRDPGLPSQOELGRSVEELWALLSPFPGGALDRVLK	1410
Db	1516	IIAKSLKNKIWNFEFRYGGFSLGVSNSQALPPSHVEVNDIAKQMKLLKLTKTDSADRFLS	1575
Qy	1411	NLTAWHSLDAQDSLKTKWFNNKGHWSVAFVNRASNAILRAHLPPGPARIHSHITTLNHP	1470
Db	1576	SLGRFMAGLDTKNNVKVFNKNGHWAISFSLNVNNAAILRANLQKGNPSOYGITAFNHP	1635
Qy	1471	LNLTKEOLSEALMASSVDVLVSTVCVFPAMSFVPASFTVLILBERVTRAKHLOLMGGLSP	1530
Db	1636	LNLTKQQLUSEVALMTTISVDVLVSI1CVLFAMSFVPASFVFLIQRBSRKAHLQIFSGVKP	1695
Qy	1531	TLYWLGNFLWDMCNLYVPACIVVLI1FLAFQORAYVAPANLIPALLLLLLLLGWSTITLMPY	1590

Db 1696 VIVWLSNFWMDNCNVVYPATLVIIIFCFQOKSFSVSTNLVPVALLLLLLGWSITPLMYP 1755
 Qy 1591 ASPPFFSVPTAYVVLTCINLFIGINGSMTAFVLELFSDDQKLOEVSRLIKOVFLIPFHFCL 1650
 Db 1756 ASVFVKIPTSAYVVLTSVNLFIGINGSVAFTVLELFTNNKLNDINDILKSQVLIFFHFCL 1815
 Qy 1651 GRGLIDMVNRNOAMADAPERLGDGQPOPLRWEVVVGKNLLAMVTQGPFLLFTLLLQHRSQ 1710
 Db 1816 GRGLIDMVNQKQAMADALERFGNERFVSPLSWDLVGRNLFAMAVEGVVFFITVLIOYRF 1875
 Qy 1711 LLQPVRVSRPLICEEDVEDVARERVRVOGATOGDVLVLRNLTKVYRGORMPADVRLCLG 1770
 Db 1876 IRPRPVKAKLPPLNDEDEDVRRQRILTDCGGQNDDILEIKELTKIYRKRPAPVDRICTG 1935
 Qy 1771 IPGECEGLLGVNGAGKTSTFRVMVTDTLASRGEAVLAGHSVAREPSPAHLMSGCYCPSD 1830
 Db 1936 IPGECEGLLGVNGAGKSTTFKMLTGDTPTVRGDPAFLNKNSILSNIEHVQNMGYCQPD 1995
 Qy 1831 AIFELLTGRHEHLELARLGRVPEAQVAQTAGSGLARLGLSWADRDPAGTVSGGNKRKLAT 1890
 Db 1996 AITELLTGREHVEFFALLRGPKEVGFGEWAIRKGLGVKYGEKVASNTSGGNKRKLST 2055
 Qy 1891 ALALVGPDAVVFDELPTTGMDSARPRLWNLSLLAVVREGESVMLTSHSMEECALCSRLA 1950
 Db 2056 AMALIGPPPVPVDELPTTGMDSARPRLWNLSLLAVVREGESVMLTSHSMEECALCTRMA 2115
 Qy 1951 IMVNGRFRCLGSPOHLKGREAAHGHTLTRVPAARS--QAAAAFVAABFPQSELREAHHGR 2008
 Db 2116 IMVNGRFRCLGSQVHLKNRGDGTYIVRTAGSNPDLKPVOEPFGLAFFGSVLKEKHNM 2175
 Qy 2009 LRFQLPGGCARALARFEGELAVHGAHGVDFSVSQTMLBEVFLYFSKQCKDE---DTE 2065
 Db 2176 LOYQL-PSSLSLARIFSILQSQKKRLHIEDYSVSQTTLDQVFVNFPAKDQSDDDLKDL 2234
 Qy 2066 EQKEAGVGVDPAQGLQHPKRVSQFDLDPSTAET 2098
 Db 2235 LHKNQTV-VDVAV-----LTSFLQDEKVKES 2259

RESULT 9
 ABC1_HUMAN STANDARD; PRT; 2261 AA.
 ID AC Q95477; Q96S86; Q9GT85; Q9NQV4; Q9UN06; Q9UN07; Q9UN08; Q9UN09;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE ATP-binding cassette, sub-family A, member 1 (ATP-binding cassette transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux regulatory protein).
 GN Name=ABCAL; Synonyms=ABC1, CERP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 NX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20345099; PubMed=1088428;
 RA Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y., Freeman-Lia, Cheng J.-F., Osorio J., Remaley A.T., Yang X.-P., Haendenschild C.C., Prades C., Chimini G.E., Blackmon E.B., Francois D.S., Duverger N., Rubin E.M., Rosier M., Denefle P., Fredrickson D.S., Brewer H.B. Jr.;
 RA "Complete genomic sequence of the human ABCAL gene: analysis of the human and mouse ATP-binding cassette A promoter";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Skin;
 RC Schwartz K., Lawn R.M., Wade D.P.;
 RA "ABCAL gene expression and apoA-I-mediated cholesterol efflux are regulated by LXr.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

- [3] SEQUENCE FROM N.A.
RX MEDLINE=21251004; PubMed=11352567; DOI=10.1006/geno.2000.6467;
RA Qiu Y., Cavellier L., Chiu S., Yang X., Rubin E., Cheng J.-F.;
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies
revealing novel regulatory sequences.";
RL Genomics 73:66-76(2001).
[4]
RN SEQUENCE FROM N.A.
RA Tanaka A.R., Abe-Dohmae S., Arakawa R., Sadanami K., Kidera A.,
RA Kioka N., Amachi T., Yokoyama S., Ueda K.;
RT "A new topological model of functional human ABCA1-signal peptide
cleavage and glycosylation of a large extracellular domain.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99194549; PubMed=10092505; DOI=10.1006/bbr.1999.0406;
RA Langmann T., Klucken J., Reil M., Liebisch G., Luciani M.F.,
RA Chimini G., Kaminski W.E., Schmitz G.;
RT "Molecular cloning of the human ATP-binding cassette transporter 1
(hABCI): evidence for sterol-dependent regulation in macrophages.";
RL Biochem. Biophys. Res. Commun. 257:29-33(1999).
[6]
RN SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99364413; PubMed=10431238;
RA Rust S., Rosier M., Funke H., Real J., Amoura Z., Piette J.-C.,
RA Deleuze J.-F., Brewer H.B., Duverger N., Deneffe P., Assmann G.;
RT "Tangier disease is caused by mutations in the gene encoding ATP-
binding cassette transporter 1.";
RL Nat. Genet. 22:352-355(1999).
[7]
RN PHOSPHORYLATION SITES SER-1042 AND SER-2054.
RX MEDLINE=22289331; PubMed=12196520; DOI=10.1074/jbc.M204923200;
RA See R.H., Caday-Malcolm R.A., Singaraja R.R., Zhou S., Silverston A.,
RA Huber M.T., Moran J., James E.R., Janoo R., Savill J.M., Rigot V.,
RA Zhang L.H., Wang M., Chalmi G., Wellington C.L., Tafuri S.R.,
RA Hayden M.R.;
RT "Protein kinase A site-specific phosphorylation regulates ATP-binding
cassette A1 (ABCA1)-mediated phospholipid efflux.";
RL J. Biol. Chem. 277:41835-41842(2002).
[8]
RN REPRESSION BY ZNF202.
RX MEDLINE=21192304; PubMed=11279031; DOI=10.1074/jbc.M100218200;
RA Porach-Oezueruemez M., Langmann T., Heimerl S., Borsukova H.,
RA Kaminski W.E., Drobnik W., Honer C., Schumacher C., Schmitz G.;
RT "The zinc finger protein 202 (ZNF202) is a transcriptional repressor
of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene
expression and a modulator of cellular lipid efflux.";
RL J. Biol. Chem. 276:12427-12433(2001).
[9]
RN INDUCTION BY LIPOPOLYSACCHARIDE.
RX PubMed=12032171;
RA Kaplan R., Gan X., Menke J.G., Wright S.D., Cai T.-Q.;
RT "Bacterial lipopolysaccharide induces expression of ABCA1 but not
ABCG1 via an LXR-independent pathway.";
RL J. Lipid Res. 43:952-959(2002).
[10]
RN REVIEW ON VARIANTS.
RX MEDLINE=22790907; PubMed=12763760;
RA DOI=10.1161/01.ATV.0000078520.99539.77;
RA Singaraja R.R., Brunham L.R., Vissecher H., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Efflux and atherosclerosis: the clinical and biochemical impact of
variations in the ABCA1 gene.";
RL Arterioscler. Thromb. Vasc. Biol. 23:1322-1332(2003).
[11]
RN VARIANTS HDLD2 THR-1091 AND 1893-GLU-ASP-1894 DEL.
RX MEDLINE=20001430; PubMed=10533863;
RA Marcil M., Brooks-Wilson A., Clee S.M., Roomp K., Zhang L.-H., Yu L.,
RA Collins J.A., van Dam M., Molhuizen H.O.F., Loubser O.,
RA Ouellette B.F.F., Senen C.W., Fichter K., Mott S., Denis M.,
RA Boucher B., Pimstone S., Genest J. Jr., Kastelein J.J.P., Hayden M.R.;
RT "Mutations in the ABC1 gene in familial HDL deficiency with defective
cholesterol efflux.";
RL Lancet 354:1341-1346(1999).
[12]
RN VARIANTS HDLD1 ARG-597 AND ARG-1477, AND VARIANT HDLD2 LEU-693 DEL.
RX MEDLINE=99364411; PubMed=10431236;
RA Brooks-Wilson A., Marcil M., Clee S.M., Zhang L.-H., Roomp K.,
RA van Dam M., Yu L., Brewer C., Collins J.A., Molhuizen H.O.F.,
RA Loubser O., Ouellette B.F.F., Fichter K., Ashbourne-Excoffon K.J.D.,
RA Senen C.W., Scherer S., Mott S., Denis M., Martindale D.,
RA Frohlich J., Morgan K., Koop B., Pimstone S., Kastelein J.J.P.,
RA Hayden M.R.;
RT "Mutations in ABC1 in Tangier disease and familial high-density
lipoprotein deficiency.";
RL Nat. Genet. 22:336-345(1999).
[13]
RN VARIANTS HDLD1 SER-590; SER-935 AND VAL-937, AND VARIANTS ALA-399 AND
MET-883.
RX MEDLINE=99364412; PubMed=10431237;
RA Bodzioch M., Orso E., Klucken J., Langmann T., Boettcher A.,
RA Diederich M., Drobnik W., Barlage S., Buechler C.,
RA Porach-Oezueruemez M., Kaminski W.E., Hammann H.W., Oette K.,
RA Rothe G., Aslanidis C., Lackner K.J., Schmitz G.;
RT "The gene encoding ATP-binding cassette transporter 1 is mutated in
Tangier disease.";
RL Nat. Genet. 22:347-351(1999).
[14]
RN VARIANTS HDLD1 ILE-929; ARG-597 AND ARG-1477, AND VARIANTS HDLD2
LEU-693 DEL; THR-1091; 1893-GLU-ASP-1894 DEL AND LEU-2150.
RX MEDLINE=20540002; PubMed=11086027;
RA Clee S.M., Kastelein J.J.P., van Dam M., Marcil M., Roomp K.,
RA Zwarts K.Y., Collins J.A., Roelants R., Tamasa N., Stulc T.,
RA Suda T., Ceska R., Boucher B., Rondeau C., Desouch C.,
RA Brooks-Wilson A., Molhuizen H.O.F., Frohlich J., Genest J. Jr.,
RA Hayden M.R.;
RT "Age and residual cholesterol efflux affect HDL cholesterol levels and
coronary artery disease in ABCA1 heterozygotes.";
RL J. Clin. Invest. 106:1263-1270(2000).
[15]
RN VARIANTS HDLD1 ASN-1289 AND HIS-1800.
RX MEDLINE=20171564; PubMed=10706591;
RA Brousseau M.E., Schaefer E.J., Dupuis J., Eustace B.,
RA Van Erdevewegh P., Goldkamp A.L., Thurston L.M., FitzGerald M.G.,
RA Yasek-McKenna D., O'Neill G., Eberhart G.P., Weiffenbach B.,
RA Ordoas J.M., Freeman M.W., Brown R.H. Jr., Gu J.Z.;
RT "Novel mutations in the gene encoding ATP-binding cassette 1 in four
Tangier disease kindreds.";
RL J. Lipid Res. 41:433-441(2000).
[16]
RN VARIANT HDLD1 ASP-1046, VARIANT HDLD2 CYS-230, AND VARIANTS LYS-219;
ILE-825; MET-883 AND LYS-1587.
RX MEDLINE=20396633; PubMed=10938021;
RA Wang J., Burnett J.R., Near S., Young K., Zinman B., Hanley A.J.G.,
RA Connelly P.W., Harris S.B., Hegeler R.A.;
RT "Common and rare ABCA1 variants affecting plasma HDL cholesterol.";
RL Arterioscler. Thromb. Vasc. Biol. 20:1983-1989(2000).
[17]
RN VARIANT HDLD1 TRP-587, AND VARIANT LEU-2168.
RX MEDLINE=21157002; PubMed=11257260;
RA Bertolini S., Pisciotto L., Seri M., Cusano R., Cantafora A.,
RA Calabresi L., Franceschini G., Ravazzolo R., Calandra S.;
RT "A point mutation in ABC1 gene in a patient with severe premature
coronary heart disease and mild clinical phenotype of Tangier
disease.";
RL Atherosclerosis 154:599-605(2001).
[18]
RN VARIANTS LYS-219; MET-883 AND ASP-1172.
RX MEDLINE=21157003; PubMed=11257261;
RA Brousseau M.E., Bodzioch M., Schaefer E.J., Goldkamp A.L., Kiehl D.,
RA Probst M., Ordoas J.M., Aslanidis C., Lackner K.J.,
RA Bloomfield Rubins H., Collins D., Robins S.J., Wilson P.W.F.,
RA Schmitz G.;
RT "Common variants in the gene encoding ATP-binding cassette transporter
1 in men with low HDL cholesterol levels and coronary heart disease.";

Db 1773 VNLPIGINGSVATFVLELFTDNKLNINDILKSVFLIPPHFCGLRGLIDMVKNQAMADAL 1832
Qy 1668 ERLGDRQPOSURWVGKLLAMVIOGFLPLFTLLQHRSQLLPQPRVSLPLIGRED 1727
Db 1833 ERFGENRFVPSLSDVGNLPMAMVEGVVFLIIVLQYRFFIRPRPNVAKLSPLNDED 1892
Qy 1728 EDVABERERVQATQGDVLRNLTKVYRGQMPAVDRCLCLGIPPGECFGLGVNGAGK 1787
Db 1893 EDVRRERQILDDGGONDILEIKELTKIYRRKRKPAVDRIKVIIPGECFGLGVNGAGK 1952
Qy 1798 TSTFMVVTGDTLASGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLAR 1847
Db 1953 STSFKMLTGDITVTRGDAPLANKNSLNIHEVQNMGYCPQSDAIFELLTGREHVEFFAL 2012
Qy 1848 LRGVPEAOVAGTAGSLARGLSVADRPAGYSGNKKRLATALLVGDPAVFLDDEPT 1907
Db 2013 LRGVPEKEVGKGEWAIRKGLVTKGEKAGYSGNKKRLSTAWALIGGPPVFLDDEPT 2072
Qy 1908 TGMDSARRFLWNSLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLK 1967
Db 2073 TGMDSARRFLWNSLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLK 2132
Qy 1968 GRPAAGHTILTRVPAARS--QPAALFAVAAFPSELRBAHGRLRPQLPGRCALARVF 2025
Db 2133 NRFGDYITVRIAGSNPDLKPVDFFGLAPFGSVPEKEGRNMLQYQL-PSLSLSLARIF 2191
Qy 2026 GELAVHGAHGVEDPSVQTMLEEVLYPSKQDKDE--DTEEOKEAGVGDPAAGLQH 2082
Db 2192 SILSOSKRLHIEDYSVQTLQDQVFNPAKQDQSDHHLKDLSLHKNQTV-VDVAV---- 2246
Qy 2083 PKRVSOFLDDPSTART 2098
Db 2247 ---LTSPLQDEKVKES 2259

RESULT 10
Q802B2
ID Q802B2 PRELIMINARY; PRT; 2201 AA.
AC Q802B2;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE ATP-binding cassette 1.
GN Name=Abca1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Ananthanarayanan M., Mirza M.F.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY208182; AAO53557.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 2201 AA; 246553 MW; B1472978BFC3E688 CRC64;

Query Match 50.6%; Score 5510.5; DB 2; Length 2201;
Best Local Similarity 49.5%; Pred. NO. 0;
Matches 1108; Conservative 361; Mismatches 575; Indels 195; Gaps 28;

Qy 15 LPSAGTVPWLOGLICNVNNTCPPOLTPGEEPRGLSNFNDLSVRLADARTVLGGASAH 74
Db 1 MPSAGTLPWVOGIIICNVNNTCPFRPTPGEAFGVGNFNKSI VSRLPFSQAQRILLYSQKDT 60
Qy 75 TLAGLGLKIATLRAAR-----STAQPQPTKQSPLE----- 104
Db 61 SIRDHMKVLRTLQIQIKHNSNKLQDPLVDNETTSGFLQHSLSLPSAVDNLLQADVSLQ 120
Qy 105 -----PPMLDVAEILL----- 114
Db 121 KVFLOGVOLHSLASLNGSKLEBEIRPEDLVKVSALCSLPREKLDAPERELRSNMDILKPMV 180
Qy 115 -----TSLRTESGLGALQAQEPHLSLEAAEDLAQELLALRLSLVELRALILQ-----RP 164
Db 181 TKLNTSLTTPQHLLA-----EATTLILDSLGLAQELFTKSDWDRQEVMTNTVNN 233
Qy 165 RGTSGPL-ELLSEALCSVRGSPSTVGPSNLWYASDLMLVGV----QBPESALPDSSASP 219
Db 234 SGSSTQIYQAVSRIVCGHPEGGLKIKSLNWDNNYKALFGGNGTEEDTDTFYDNSTTP 293
Qy 220 ACSELIGALDSHPRLSLRLWRRLKPLILGLKLLFADPTPTKRLMAQVNTFPELTLRLDVR 279
Db 294 YCNLDMKNLESPLSRIRIWKALKPLILGLKLLFTPTPATRQVMAEVNKTQSLALFPDL 353
Qy 280 EVWEMLGPRIFTFWMNDSNVAMLQRLQMOQ--DEGRORPPGGRDHMEALRSFL----- 331
Db 354 GMEELSPQIWTFWESSQEMDLVRPMLDLRNGDFWERKLDGLYTAQDIWAFKLNPE 413
Qy 332 --DFSGGYSQMDAHADVGLVGTIGRVTESLSDKLEAAPSEALVSRALQLLAEHRFW 389
Db 414 VQSPNGSVYTWREAFNETNQAIQITISRFMECVNLNKLPIPTVTLINKSMDDLARKFS 473
Qy 390 AGVVPGLGEDSSDTEHTPTDLG--PGHVRIRKIMDIDVVTTRTKIRDRFNDPGAPADPL 447
Db 474 AGIDFTG-----ITPDSVELPHHVKDKIMDIDNVERTNKIKGYNDPGPRADPF 523
Qy 448 TDLRYVMGGFVYVLODLVERAAVRVLSGANPRAGYVLOQMPYPCYVDDVFLVLSRLPLF 507
Db 524 EDMRYVMGGFAYLQDVVEQAIIRVLTGTEKTYVQVQMPYPCYVDDIPLRVMSRMPLF 583
Qy 508 LTLAWIYSVTITVKAIVVREKETRLRDTMRAMGLSRAVLWGLWFLSCGLGPFLLSALV 567
Db 584 MTLAWIYSVAVI IKSIIVVEKARLKETMRIMGLDNGILWFSWFISSLPLPLVSAGLVII 643
Qy 568 LKLGDIPLYPHGPVVFLLAFAVATVTSQSLLSAPFSRANLAAACGLAYPSLYLPYL 627
Db 644 LKLGDIPLYPSPSVVFLSVFAVVTILQCFILSTLFSRXNLAACGIIYFTLLYLPYL 703
Qy 628 CVAWRDLRPLAGGRVAASLLSPVAFGCGESLALLEEQEGAGQWNVGTRPT-ADVFSLAQ 686
Db 704 CVAQDVTGVSIFKIFASLLSPVAFGCGEYFALFEQOIGVQWNLFPKSVPEEDGFNLT 763
Qy 687 VSGLLLLDAALYGLATWYLEAVCPGQYGIPEPMNPFRRRSYWCGRPSPKSPAPCTPLD- 745
Db 764 SVSNMMLFDTFIYGVMTWYIEAVFPQYGIIPRWYFCTKSYWFGEBIDEKSHPGSSOKGA 823
Qy 746 PKVLVEAPGLSPGVSVRSLEKRPSPQPALRGLSLDFYQGHITAFPLHNGAGKTTTL 805
Db 824 SEICMEESEPTHLKLGVSQNLVKYVRDGMKVAVDGLALNPYEGQITSPFLHNGAGKTTM 883
Qy 806 SILSGLPPSGGSAPILGHIDVRSMAALRPHLGVCPQVNVLPDMLTVDEHVMFYGRKGL 865
Db 884 SILTGLFPPTSGTAYIILKDIRSEWNSIRQNLGVCPQHVNLFDMLTVEEHLWFAARLKL 943
Qy 866 SAAVVGPSQDRLLQDVGL-VSKQSVQTRHLSGGMRKLSVAIAFVGGSVQVILDEPTAGV 924
Db 944 SEKHVKAEMEQLDVLGPPSKLSKTSQLSGGMRKLSVALA FVGGSKVILDEPTAGV 1003
Qy 925 DPASRRGIWELLKYRGRRTILSTHLLDEALLGDRVAVVAGGRGCCGSPFLRRHLG 984
Db 1004 DPYRRGIWELLKYRGRRTILSTHLLDEADILGDIRIAITISHGKLCVCGSSFLKNOLG 1063

Db 346 KAFGLIDSTRKDPYISYDKRTTTPCNALIOSLESNPLTKIAWRAAKPLVMGKILPTDPS 405
Qy 257 FTRKMAQVNRTEFEELTLRLDREVMELGPRIFTFMDSNVAMLORL----- 306
Db 406 AVRRILONANSTPEELERLKLVAEEVGPQIWPFDRTSTOMTIRDTLENPTVKGFLN 465
Qy 307 -QMDEGRQRPGRGDHNEALRSFLDPG-----SGYSWQDAHADVGLVGTGRV 357
Db 466 SOLGBEGIT-----AEAMNLFHLKGPRESQADDMANFDWRDVFNITDRTLTSKY 516
Qy 358 TECLSLDKLEAPSEAAVSRALQALLAHRFWAGVVLGPESSDPTHEPTDLGPHVR 417
Db 517 LECLILDKPESYDDEIQLTORALSLEENRFAGVVF-----PMYPTWSALPTHVK 568
Qy 418 IKIRMDIDVTRTKIRDFWDPGAADPLTDLRYVMGFFVYLQDLVERAAVRVLSGAMP 477
Db 569 YKIRMDIDVVEKTKIKORYWDSGRADPVEDFRVIWGGPAYLQDMIEGITRSQAQVEV 628
Qy 478 RAGLYLOMPYPCYVDDVFLVLSRSLPLFLTLAMIYSVTLTVKAVVREKETRLRDTMRA 537
Db 629 PVGIYLOMPYPCFVDDSFMIILNRCFPIFMVLAMIYSVMTVKISIVLEKELRLKXELKN 688
Qy 538 MGLSAVLWGLFSLCLGPPLSAALLVLVLKGLDILPYSHGVVPLFLAFAVATVQS 597
Db 689 QGVSNVAVICTWFLDSFSTMSMSIFLLTIFIMHGRILHYSNPFILFLPLAFSTATIMQC 748
Qy 598 FLLSAFFSRANLAAACGLAYPSLYPLVYLCVWDRDLRPAAGRVAASLLSPVAFGCGES 657
Db 749 FLLSTFPSEASLAAACSGVIYTLYLPHILCFANQDRMTADLKMVLSLLSPVAFGCTEY 808
Qy 658 LALLEQGEAGQWNVGTRP-TADVFLSAQVSLGLLLDAAALYGLATWYLEAVCPQYQYIP 716
Db 809 LARPEQGLGLQWSNIGKSPMGDEDFSLMSKMWMLDAAALYGLLAWYLDQVFPNGYGP 868
Qy 717 EPNWPFPRSYCG-----PRPKSPACPTPL-DPK-----VLVEAPPLSGP 760
Db 869 LPWYFLQESYVLGEGCSTREERALEKTEPTTEEMEDPEHPGINDAPFERELGLVPG 928
Qy 761 VSVRLERKPPGSPALRGLSLDFVQGHITAFGLHNGAGKTTLSILSGLPSPGGSF 820
Db 929 VCVKLVKIPYSPVADRLNITYENQITAFGLHNGAGKTTLSILGTLPPSTGTVL 988
Qy 821 ILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTVDHWFYGRUKGLSAAVGPQDRLLQD 880
Db 989 IGGKDIETSLDVRQSLGMCPOYNILFPHILTVAEHILFYAQLKGKSWEEAQLMEAMLED 1048
Qy 881 VGLVSKQSVOTHLGGHQKLSVAIAFVGGQVILDEPTAGVDPASRGITWELLKYR 940
Db 1049 TGLHHRNEEAQDLSGHQKLSVAIAFVGGAKVILDEPTSGVDPPYRRSITWDLKKYR 1108
Qy 941 EORTLILSTHLDREALLGDRVAVVAGGRGCCGSPFLRRLHLSGGYLLTVKARLPLTT 1000
Db 1109 SGRTIIMSTHMDREALLGDRIRIISQGLYCSGTFPLKNCFTGPTLTVR-----KMN 1165
Qy 1001 NEKADTDMEGSVDTREKKNQSGSRVG--TP-----QLLALVQHVWPGARLVEEL 1049
Db 1166 IQSQTGCGTSCCASEGFSTRCPAHIDEITPEQVLGDGVNLMQVHHPVEAKLVECI 1225
Qy 1050 PHELVLVLPYTGADGSAFATLRELDTRLAELRLTYGIGISDTSLEEIFLKVVESCAADTD 1109
Db 1226 GQELIFLFPKNPKQKQVASFRELEDTLADLGLSSFGISDTPLEEIFLKVT----- 1277
Qy 1110 MEDSGQHLCGTGIAGLDVTLRLK-MPQETALE-----NG---EPAGSAPETDQSGPD 1160
Db 1278 -EDSUGGLFAGGTQCKENLRHPWLSPEKARQIPQSGNCGSXEPAPHPBQPSSEPE 1335
Qy 1161 AVGRVQ-GWALTQOQLOALLKRRFLARRSRRLGPAQIVLPALFVGLALVFSLIIVPPFGH 1219
Db 1336 ARSRLNTGAQLIIHQVALLVKKRFHTIRSHKDFLAQIVLPATFVPLALMLSLIIVPPFGE 1395
Qy 1220 YPALRLSPTMYGAQVFSFSDAPGPGRARLLEALLQERAG-----LEPPVQHSS 1269
Db 1396 YPALILHPWYQQYTFFSLDQPSGEQLAALADVLNKPFGNRCCLKBEWLPEYPCGNST 1455

Qy 1270 HRFSAPEVPAEYAKVLASGNWTPESPACQSQCPGARRLLPDCPAAAGPPPPQAVTGS 1329
Db 1456 -PWKTPSVSPNITHLFQKQWTPKPSPCRSCTREKLTMLPECEGAGGLPPQRIORS 1514
Qy 1330 GEVQNLTRNLSDPLVITYPRLVRQGLKTKKQVNEVYGGSPSLGGRDGLP-SGQELQR 1388
Db 1515 TEILODLTRNISDPLVITYPALIRSSLSKFWNEQRYGGISIGKLPILPITGEALVE 1574
Qy 1389 SVEELWALLSLPGGALDR-VLKNLTANAHSIDAODSLKIFWNNKGWMSVAFVNRASNA 1447
Db 1575 PLSHLQGMN-VSGGPITREASKEMPAFLKHLETDNIKWPNNGWHALVSFLVAHNA 1633
Qy 1448 ILRAHLPFGPARHAHSITTLNHLPLMLTKEQLSEAAALMASSVDVLVSIICVFAMSVFPAF 1507
Db 1634 ILRTSLHKDKNPEEYGIIVISOPLNLTKEOLEITVLTASVDVAIVCIVFAMSVFPAF 1693
Qy 1508 TLVLIBERTRAKHLQMLGGLSPTLYWLNFLNMDMNCNVLVPAICIVLFLAPQQAAYVAP 1567
Db 1694 VLYLTQERVNKAKHLQFVSGVSPITYTLNLFMDIMNVAASALVVGIFVGFQKCAVTS 1753
Qy 1568 ANLPALLALLLYGWSITPLAMVPASPFPFSPSTAVVLTCTNLFTGINGSMATFVLELPS 1627
Db 1754 ENLPALIALMLYGMVAVIPMMYPASFLFDVPSTAYVALSCANLFTGINSATFLELPE 1813
Qy 1628 DQ-KLQEVSRILKQVLIFFPHFCLGRGLIDMVNRQAMADAFERLGDROFQSPLRVEVVGK 1686
Db 1814 NNWTLRFRNMLKLLIIFPHFCLGRGLIDALSQAQVTVVYARFGEESTNPFQWDLIGK 1873
Qy 1687 NLLAMVIOGFLPLFTLLQHR---SOLLPOVRSLPLLGESEDEVARERERVVQGAQ 1743
Db 1874 NLVMAABGVVLLLTFLFIQHPFLTRVWSEPAKE--PII-DEDDVAERQRIISGGNK 1930
Qy 1744 GDVLVRLNLTKVYQORMPAVDRLCLGTPCECFLGNGAGKTSTFRWVGTDLASRG 1803
Db 1931 TDILRLNELTKIYSGTSSPAVDRLCVGRPCECFLGNGAGKTTTKMLTGDTTVTSG 1990
Qy 1804 EAVLAGHVAAREPSAAHLSMGYCPOSDAIFELLTGREHLELLIARLVRGYPEAQVATAGSG 1863
Db 1991 DATAGKILTNISDVHOSMGYCPQFQDAVDDLLTGREHLYLARLVRGYPADIEIVANWS 2050
Qy 1864 LARGLSWYADRPAGTYSGGNKRKLATALALVGPVAVFLDEPTTGMPPSARRPLWNSLL 1923
Db 2051 IQSLGLSLYADRLVGTYSGGNKRKLSTALALMGCPPLVLLDEPTTGMPPARRMLWNTIV 2110
Qy 1924 AVVREGSVMLTSHSMECEALCSRLATMNGFRCLGSPHKLGRFAAGHTLTLRVPA 1983
Db 2111 SIIREGRAVLTSHSMECEALCTRLAIMVXGTFOCLGTIOHLKYKFGDGYIVTMKIKSP 2170
Qy 1984 RS-----QPAFAAFVAAAPPSELREAHGGRRLRFPOLPGRCALARVFGELAVHGAHGV 2037
Db 2171 KEDLLPDLNPFVEQFQGNFSGVQERHYNNMLQFV---SSSLARIPELLISHKDSLLI 2227
Qy 2038 EDFSQVQMLBEVFLYFSKQDQKDED 2063
Db 2228 EYSVTQTTLDQVFNFAKQQTETHD 2253
RESULT 12
AAAR87835 ID AAR87835 PRELIMINARY; PRT; 2269 AA.
AC AAR87835;
DT 20-MAY-2004 (TremBLrel. 27, Created)
DT 20-MAY-2004 (TremBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TremBLrel. 27, Last annotation update)
DE ABCA4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS23;

[illegible]

2228 EEYSVTQTTLDQVFNFAKQQTETHD 2253

RESULT 14
 AAR87836 PRELIMINARY; PRT; 2269 AA.

ID AAR87836 AC AAR87836; PRT; 2269 AA.
 AC AAR87836; PRT; 2269 AA.
 DT 20-MAY-2004 (TrEMBLrel. 27, Created)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE ABCA4.

OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=P1117;
 RC PubMed=15064680;
 RX Kijas J.W., Zangerl B., Miller B., Nelson J., Kirkness E.F.,
 RA Aguirre G.D., Acland G.M.;
 RT "Cloning of the canine ABCA4 gene and evaluation in canine cone-rod
 RA dystrophies and progressive retinal atrophies.";
 RT Mol. Vision 10:223-232 (2004).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=P1117;
 RC Kijas J., Zangerl B., Miller B., Nelson J., Kirkness E., Aguirre G.,
 RA Acland G.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY427779; AAR87836.1; -;
 SQ SEQUENCE 2269 AA; 256491 MW; 9B6E349FF17895A1 CRC64;

Query Match 45.0%; Score 4903; DB 2; Length 2269;
 Best Local Similarity 45.5%; Pred. No. 2.3e-285;
 Matches 1023; Conservative 347; Mismatches 654; Indels 222; Gaps 37;

Qy 1 PPLEHHECHFPNKPPLPSAGTVPMQLGLICNNVNTCPQLTPGEPGRLSNFNSLSRLL 60
 Db 47 PLSQHECHFPNKAMPAGMLPMLQGMFCNNVNPFCQNPPTGPSPGVSNYNNLSILARVF 106
 Qy 61 ADARTVLGGASAHRTLAGLGLKLTATLRAASTAQPPQTKOSPLEPPMLDVAE----- 112
 Db 107 RDPQELLDADPQRQHFQHWKBEQTLSRLMDTLRTHPVRAGIRDIRDLKDEETLTF 166
 Qy 113 -----LLTSLARTSL----- 123
 Db 167 LMKNIGLSDSVVYLLNSQVRPEQFAHGVPDMLMKDIACSETLLERFIIFSQRGAQTVR 226
 Qy 124 -----GLALQAQ-----EPLHSLEAEADLAQELLARS-----LVELRA 159
 Db 227 DAMCSLSQGLQWVEDLYANVPDFKLFRVLPTLLDSSQGINLRSGWRVFSDISSIRE 286
 Qy 160 LLQRP-----RGTSGP-----LELLSEALCSV-RGPSSTVGPSLNWYEASDL 200
 Db 287 FIHRPSVEDLLWVTKPTQTGGPETAFLQMSILSDLLCGYPEGGSRVF-SFNWYEDNNY 345
 Qy 201 MELVGQEPESALP-----DSSLSPACSELI GALDSHPLSRLLRRLKPLILGLKLLFADPTP 256
 Db 346 KAFGLIDSTRKDPIYSYDKRTTTCNALIQSLSENPLTKIAWRAAKPLVMGKILFTPDSP 405
 Qy 257 FTRKLMQAVNRTFEELTLRADRVREVMELGPRTFENMDSNNVAMLQRL----- 306
 Db 406 AVRRILQNAVNSTFEELERKLKVAMSEVGPQIWFYFDRSTQMTIMHIDTLENPTVKGFLN 465
 Qy 307 -QMDEGRRQPRPGGRDHMEALRSFLDPG-----SGGYSQDAHADVGHVLVGLTGRV 357
 Db 466 SOLGEEGIT-----AEAMLNFKHGPRESQADDMANFDRVDVENITDRTLRLTSKY 516
 Qy 358 TECLUSDKLFAA PSEALVSRALQLLAHRRFQWAGVFLGPEDSDDTPEHTTDLPGPHVR 417
 Db 517 LECILILKFFSYDDEIQLTQRALSLEENRFWAGVF-----PMYPTWTSALPTHVK 568

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:28:54 ; Search time 106.335 Seconds
(without alignments)
7104.204 Million cell updates/sec

Title: US-09-995-542-6

Perfect score: 10896

Sequence: 1 PPLEHHECHFPNKPPLPSAGT.....QHPKRVSQFLDDPSTAETVL 2100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10896	100.0	2100	9	US-09-995-542-6
2	10896	100.0	2146	9	US-09-995-542-5
3	10883	99.9	2146	15	US-10-182-006-2
4	10883	99.9	2180	15	US-10-332-447-27
5	10880	99.9	2144	9	US-09-858-194-2
6	10880	99.9	2144	14	US-10-154-419-2
7	10874.5	99.8	2147	17	US-10-618-281-44
8	10392.5	95.4	2059	15	US-10-114-270-176
9	9213.5	84.6	1873	15	US-10-182-006-4
10	8389.5	77.0	2121	9	US-09-995-542-3
11	8389.5	77.0	2167	9	US-09-995-542-2
12	7596	69.7	1550	9	US-09-995-542-8
13	5610.5	51.5	2261	14	US-10-313-641-9

14	5610.5	51.5	2261	14	US-10-313-641-10	Sequence 10, Appl
15	5610.5	51.5	2261	14	US-10-428-551-9	Sequence 9, Appl
16	5610.5	51.5	2261	14	US-10-428-551-10	Sequence 10, Appl
17	5610.5	51.5	2261	15	US-10-452-510-1	Sequence 1, Appl
18	5610.5	51.5	2261	15	US-10-617-334-1	Sequence 1, Appl
19	5610.5	51.5	2261	16	US-10-745-377-5	Sequence 5, Appl
20	5610.5	51.5	2261	16	US-10-744-465-1	Sequence 1, Appl
21	5610.5	51.5	2261	16	US-10-833-679-1	Sequence 1, Appl
22	5610.5	51.5	2261	17	US-10-872-113-5	Sequence 5, Appl
23	5603.5	51.4	2261	14	US-10-340-097-118	Sequence 118, App
24	5603.5	51.4	2261	14	US-10-336-215-118	Sequence 118, App
25	5603.5	51.4	2261	14	US-10-336-219-118	Sequence 118, App
26	5594.5	51.3	2261	10	US-09-984-827-128	Sequence 128, App
27	5594.5	51.3	2261	10	US-09-984-827-132	Sequence 132, App
28	5594.5	51.3	2261	10	US-09-984-827-134	Sequence 134, App
29	5594.5	51.3	2261	10	US-09-984-827-136	Sequence 136, App
30	5593.5	51.3	2261	9	US-09-995-542-11	Sequence 11, Appl
31	5593.5	51.3	2261	9	US-09-984-827-2	Sequence 2, Appl
32	5593.5	51.3	2261	10	US-09-984-827-127	Sequence 127, App
33	5593.5	51.3	2261	10	US-09-984-827-129	Sequence 129, App
34	5593.5	51.3	2261	10	US-09-984-827-130	Sequence 130, App
35	5593.5	51.3	2261	10	US-09-984-827-131	Sequence 131, App
36	5593.5	51.3	2261	10	US-09-984-827-133	Sequence 133, App
37	5593.5	51.3	2261	10	US-09-984-827-135	Sequence 135, App
38	5588.5	51.3	2261	10	US-09-984-827-133	Sequence 133, App
39	5586.5	51.3	2263	15	US-10-276-774-2326	Sequence 2326, Ap
40	5531.5	50.8	2201	9	US-09-995-542-9	Sequence 9, Appl
41	5522.5	50.7	2201	14	US-10-170-385-293	Sequence 293, App
42	5522.5	50.7	2201	14	US-10-331-496A-29	Sequence 29, Appl
43	5522.5	50.7	2201	15	US-10-429-160-4	Sequence 4, Appl
44	5522.5	50.7	2201	16	US-10-465-498-139	Sequence 139, App
45	5470.5	50.2	2233	16	US-10-465-498-140	Sequence 140, App

ALIGNMENTS

RESULT 1

US-09-995-542-6
; Sequence 6, Application US/09995542
; Patent No. US2002012647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-6

Query Match	100.0%;	Score 10896;	DB 9;	Length 2100;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2100;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	PPLEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCFPOLTGEGPGRLSNFDLSVRL 60		
Db	1	PPLEHHECHFPNKPPLPSAGTVPWLOGLICNVNNTCFPOLTGEGPGRLSNFDLSVRL 60		
Qy	61	ADARTVLGGASAHRTLAGLKIATLAARSTAOPTKQSPPLPMDVAELTSLRT 120		
Db	61	ADARTVLGGASAHRTLAGLKIATLAARSTAOPTKQSPPLPMDVAELTSLRT 120		
Qy	121	ESLGALGQAEPLHSLLEAEDIQAELLALRSLVELRALQRPRTGSPGLELSEALCS 180		
Db	121	ESLGALGQAEPLHSLLEAEDIQAELLALRSLVELRALQRPRTGSPGLELSEALCS 180		

Db 121 ESIGLALGOQEPHLSLLEAAEDLAQELLALRSIVELRALLOPRGTSGLPLELLSEALCS 180
Qy 181 VRGFSSTVGPLSNWYEASDLMELVGOBPESALPDSSLPACSELIGALDHPHLSRLLWRR 240
Db 181 VRGFSSTVGPLSNWYEASDLMELVGOBPESALPDSSLPACSELIGALDHPHLSRLLWRR 240
Qy 241 LKPLILGLKLLFAPDTPTRKLMQAVNRTPEELTLRLDRVREWMELGPRIETTFMNDSSNVA 300
Db 241 LKPLILGLKLLFAPDTPTRKLMQAVNRTPEELTLRLDRVREWMELGPRIETTFMNDSSNVA 300
Qy 301 MLQRLQOMQDEGRQRPGRDHEALRSPLDPSGGYSQDADHADVGLHVLGTLGRVTEC 360
Db 301 MLQRLQOMQDEGRQRPGRDHEALRSPLDPSGGYSQDADHADVGLHVLGTLGRVTEC 360
Qy 361 LSLDKLEAPSEALVSRAQLLAEHRFWAGVYVFLGPEDESSDTEHPTDPLDGPCHVRIKI 420
Db 361 LSLDKLEAPSEALVSRAQLLAEHRFWAGVYVFLGPEDESSDTEHPTDPLDGPCHVRIKI 420
Qy 421 RMDIDVVRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 480
Db 421 RMDIDVVRTNKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 480
Qy 481 LYLQOMPYPCVDDVFLVRSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL 540
Db 481 LYLQOMPYPCVDDVFLVRSRSLPLFLTLAWIYSVTLTVKAVVREKETRLRDTMRAMGL 540
Qy 541 SRAVLWLGWFLSCLGPELLSAALLVLVLKLGDLIPYSHPGVWFLFLAFAFAVAVTQSFL 600
Db 541 SRAVLWLGWFLSCLGPELLSAALLVLVLKLGDLIPYSHPGVWFLFLAFAFAVAVTQSFL 600
Qy 601 SAFFSRANLAAACGLAYFSLYLPYVLCVAVWRDLPAAGRVAASLLSPVAFPGCSLAL 660
Db 601 SAFFSRANLAAACGLAYFSLYLPYVLCVAVWRDLPAAGRVAASLLSPVAFPGCSLAL 660
Qy 661 LESQGEQAQHNHVGTRPTADVSLAQSGLLLDAAALYGLATWYLEAVCPQGVIGPEPN 720
Db 661 LESQGEQAQHNHVGTRPTADVSLAQSGLLLDAAALYGLATWYLEAVCPQGVIGPEPN 720
Qy 721 FPFRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 721 FPFRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
Qy 781 LSLDFYQCHITAFILGHNGAGKTITLSLSGLFPSPGSAFILGHDVRSMAAIRPHLGV 840
Db 781 LSLDFYQCHITAFILGHNGAGKTITLSLSGLFPSPGSAFILGHDVRSMAAIRPHLGV 840
Qy 841 PQYNVLFDMLTVDBHWFYGRKLGSAVVGPEQDRLLQDVLGVSQSVOTRHLSGQMOR 900
Db 841 PQYNVLFDMLTVDBHWFYGRKLGSAVVGPEQDRLLQDVLGVSQSVOTRHLSGQMOR 900
Qy 901 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
Db 901 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDEAELLGD 960
Qy 961 RVAVVAGRLCCCGSPFLFRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROBKN 1020
Db 961 RVAVVAGRLCCCGSPFLFRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROBKN 1020
Qy 1021 GSQSRVGTPLLALVOHWPVGARLVEPHELVLVLPYTGADHGSFATLIFRELDTRLAE 1080
Db 1021 GSQSRVGTPLLALVOHWPVGARLVEPHELVLVLPYTGADHGSFATLIFRELDTRLAE 1080
Qy 1081 LRLTYGISTLSLEEILFKVVEECAADTDMEDGSCGQHLCTGIAGLDTLRLKMPQETA 1140
Db 1081 LRLTYGISTLSLEEILFKVVEECAADTDMEDGSCGQHLCTGIAGLDTLRLKMPQETA 1140
Qy 1141 LENGEPAAGSPETDQSGPDVAVRQVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
Db 1141 LENGEPAAGSPETDQSGPDVAVRQVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
Qy 1201 ALFVGLALVFSLLIVPPFGHYPALRLSPTMYGAQVSYFFSEADPGPGRARLLEALLQAGL 1260
Db 1201 ALFVGLALVFSLLIVPPFGHYPALRLSPTMYGAQVSYFFSEADPGPGRARLLEALLQAGL 1260

Qy 1261 EBPVVOHSSHRSPAPVPAEAVKVLASGNWTPESPSPACQOCQOPGARRLLPDCPAAAGGP 1320
Db 1261 EBPVVOHSSHRSPAPVPAEAVKVLASGNWTPESPSPACQOCQOPGARRLLPDCPAAAGGP 1320
Qy 1321 PPPQAVTSGGEVVQNTLTGRNLSDFLVKTYPLRVQGLKTKKWVNEVRYGGFSLGGDRDGL 1380
Db 1321 PPPQAVTSGGEVVQNTLTGRNLSDFLVKTYPLRVQGLKTKKWVNEVRYGGFSLGGDRDGL 1380
Qy 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSDAQDSLKIWFNKNKGWHSVAF 1440
Db 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTLAWAHSDAQDSLKIWFNKNKGWHSVAF 1440
Qy 1441 VNRASNAIIRAHLPPOPPARHAHSITTLNHPNLNITKEOLSEALMASSVDVLVSI CVVFAM 1500
Db 1441 VNRASNAIIRAHLPPOPPARHAHSITTLNHPNLNITKEOLSEALMASSVDVLVSI CVVFAM 1500
Qy 1501 SFVPASFTLVLTIERVTRAKHLQMGGLSPTLYWLGNFLWDMCNLYLVPAACIVVLI FLAQ 1560
Db 1501 SFVPASFTLVLTIERVTRAKHLQMGGLSPTLYWLGNFLWDMCNLYLVPAACIVVLI FLAQ 1560
Qy 1561 QRAYVAPANLPALLLLLLLYGMSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMAT 1620
Db 1561 QRAYVAPANLPALLLLLLLYGMSITPLMYPASFFSPSTAYVVLTCINLFIGINGSMAT 1620
Qy 1621 FVLELFSQDKLQBSVRLKQVFLI PFHFCILGRGLIDMVRNOAMADAFERLGDQFQSPPLR 1680
Db 1621 FVLELFSQDKLQBSVRLKQVFLI PFHFCILGRGLIDMVRNOAMADAFERLGDQFQSPPLR 1680
Qy 1681 WEVVGKNLLAMWTQGLFLLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVARERERVQ 1740
Db 1681 WEVVGKNLLAMWTQGLFLLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVARERERVQ 1740
Qy 1741 ATQGDVLVLRNLTKVYRGQMPAVDRCLGIPGECFGLLVGNAGKTSFRMVTGDTLA 1800
Db 1741 ATQGDVLVLRNLTKVYRGQMPAVDRCLGIPGECFGLLVGNAGKTSFRMVTGDTLA 1800
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELITGREHLELLARLRCVPEAQVAQTA 1860
Db 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELITGREHLELLARLRCVPEAQVAQTA 1860
Qy 1861 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFDEPTTGMDPSARREPLWN 1920
Db 1861 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFDEPTTGMDPSARREPLWN 1920
Qy 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Db 1921 SLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Qy 1981 PAARSOPAAAFVAAEPFSGELREAHGRLRFQLPGRGRCALARVFGELAVHGAHGVEDF 2040
Db 1981 PAARSOPAAAFVAAEPFSGELREAHGRLRFQLPGRGRCALARVFGELAVHGAHGVEDF 2040
Qy 2041 SVSQTMLVEEVFLYPSKQDKDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTASTVL 2100
Db 2041 SVSQTMLVEEVFLYPSKQDKDETEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTASTVL 2100

RESULT 2

US-09-995-542-5
; Sequence 5, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR FILING DATE: 2001-11-28
; CURRENT APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-5

Query Match      100.0%; Score 10896; DB 9; Length 2146;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGCEBPGRLSNFNDSLSVRL 60
DB 47 PPLEHHECHFPNKPPLPSAGTVPWLQGLICNVNNTCFPQLTGCEBPGRLSNFNDSLSVRL 106
QY 61 ADARTVLGASAHRTLAGIKLIATLRAARSTAOQPTKQSPLEPMLDVAELTSLRT 120
DB 107 ADARTVLGASAHRTLAGIKLIATLRAARSTAOQPTKQSPLEPMLDVAELTSLRT 166
QY 121 ESLGLALGQAQEPHLSLEAEADLAQELLALRSVLVELRALLQRPRTGTSGLLELSEALCS 180
DB 167 ESLGLALGQAQEPHLSLEAEADLAQELLALRSVLVELRALLQRPRTGTSGLLELSEALCS 226
QY 181 VRGSSSTVGPNSLNWYEASDLMELVGQEPESALPDSSLSPACSELI GALDSHPLSLRWR 240
DB 227 VRGSSSTVGPNSLNWYEASDLMELVGQEPESALPDSSLSPACSELI GALDSHPLSLRWR 286
QY 241 LKPLILGKLLPAPDTPFTTKLMAQVNRTEELTLRDVREVMELGPRIFTFMDNDSNVA 300
DB 287 LKPLILGKLLPAPDTPFTTKLMAQVNRTEELTLRDVREVMELGPRIFTFMDNDSNVA 346
QY 301 MLQRLLOQDEGRRPRGGRDHMEALRSFLDPGSGGYSWODAHADVGLVGTIGRVTEC 360
DB 347 MLQRLLOQDEGRRPRGGRDHMEALRSFLDPGSGGYSWODAHADVGLVGTIGRVTEC 406
QY 361 LSLDKLEAPSEAAULSVRALQLLAEBHRFWAGVVFGLGEDSDSPTBHTPDIGPGHVRKI 420
DB 407 LSLDKLEAPSEAAULSVRALQLLAEBHRFWAGVVFGLGEDSDSPTBHTPDIGPGHVRKI 466
QY 421 RMDIDVTRTNKIRDFWDPGPAADPLTLDRVVMGGFVYLQDLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTNKIRDFWDPGPAADPLTLDRVVMGGFVYLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYCYVDDVFLVLRSLPLFTLAMIYSVTLTVKAVVREKETRLRDTWRAMGL 540
DB 527 LYLOQMPYCYVDDVFLVLRSLPLFTLAMIYSVTLTVKAVVREKETRLRDTWRAMGL 586
QY 541 SRAVLWLGWFLSCLGPFLLSAAALLVLKLDGILPYSHPGVVVFLFLAAFAVATVTSFLL 600
DB 587 SRAVLWLGWFLSCLGPFLLSAAALLVLKLDGILPYSHPGVVVFLFLAAFAVATVTSFLL 646
QY 601 SAFFSRANLAAACGGGLAYFSLYPLPYVLCVAREDRLPAGRVAASLLSPVAFPGCESIAL 660
DB 647 SAFFSRANLAAACGGGLAYFSLYPLPYVLCVAREDRLPAGRVAASLLSPVAFPGCESIAL 706
QY 661 LEEQCEGAQHNVGTRPTADVPFLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWN 720
DB 707 LEEQCEGAQHNVGTRPTADVPFLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPPEWN 766
QY 721 PPFRRSYWCGRPKPSAPCPTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQALRG 780
DB 767 PPFRRSYWCGRPKPSAPCPTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQALRG 826
QY 781 LSLDPYQGHITAFIHNAGAGKTTLSILSGLPPSPGGSAFILGHVDRSSMAAIRPHLVC 840
DB 827 LSLDPYQGHITAFIHNAGAGKTTLSILSGLPPSPGGSAFILGHVDRSSMAAIRPHLVC 886
QY 841 POYNVLFDMLTVDVHWFVGLKGLSAAVGPEQDRLLQDVLSKQSVOTRHLSCGMOR 900
DB 887 POYNVLFDMLTVDVHWFVGLKGLSAAVGPEQDRLLQDVLSKQSVOTRHLSCGMOR 946
QY 901 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDEALLGD 960
DB 947 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDEALLGD 1006
QY 961 RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTROEKN 1020
DB 1007 RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTROEKN 1066
QY 1021 GSQSRVGTPLLALVQHWPFCARLVEBELVVLVLPYTGADHSGPATLPRELDTRLAE 1080
DB 1067 GSQSRVGTPLLALVQHWPFCARLVEBELVVLVLPYTGADHSGPATLPRELDTRLAE 1126
QY 1081 LRLTYGIGSDTSLBEIFLKVVBECAADTMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1140
DB 1127 LRLTYGIGSDTSLBEIFLKVVBECAADTMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1186
QY 1141 LENGEPAAGSAPETDQSGPDVAVRQGVWALTRQQLQALLKRFLLARRSRGLFAQIVLP 1200
DB 1187 LENGEPAAGSAPETDQSGPDVAVRQGVWALTRQQLQALLKRFLLARRSRGLFAQIVLP 1246
QY 1201 ALFVGLALVFSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDFGRALLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSLIIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGDFGRALLLEALLQEAGL 1306
QY 1261 EEPVQVHSHRPSAPAEVPAEVAKVILASGNWTPESPSPACQSQPCARLLPDCPAAGP 1320
DB 1307 EEPVQVHSHRPSAPAEVPAEVAKVILASGNWTPESPSPACQSQPCARLLPDCPAAGP 1366
QY 1321 PPQAVTSGSGEVQNLGTGRNLSDFLVKTYPRLVQGLTKKWNNEVRVYGGFSLGGRDGL 1380
DB 1367 PPQAVTSGSGEVQNLGTGRNLSDFLVKTYPRLVQGLTKKWNNEVRVYGGFSLGGRDGL 1426
QY 1381 PSGELGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGHWSMAF 1440
DB 1427 PSGELGRSVEELWALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFWNNKGHWSMAF 1486
QY 1441 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSVICVFPAM 1500
DB 1487 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSVICVFPAM 1546
QY 1501 SFVPASFTLVLIERRVTRAKHLQLMGGISPTLYWLNFLWDMCNVLYVPACIVLILFLAFQ 1560
DB 1547 SFVPASFTLVLIERRVTRAKHLQLMGGISPTLYWLNFLWDMCNVLYVPACIVLILFLAFQ 1606
QY 1561 QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT 1620
DB 1607 QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT 1666
QY 1621 FVLELFSQKLQEVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROQSPULR 1680
DB 1667 FVLELFSQKLQEVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROQSPULR 1726
QY 1681 WEVVGKNLLAMVIOGPLELFTLLLOHRSQLLPOPRVRSPLPLGDEEDVARERERVQ 1740
DB 1727 WEVVGKNLLAMVIOGPLELFTLLLOHRSQLLPOPRVRSPLPLGDEEDVARERERVQ 1786
QY 1741 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCEFCGLLVNGVAGAKTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCEFCGLLVNGVAGAKTSTFRMVTGDTLA 1846
QY 1801 SRGSAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVOTA 1860
DB 1847 SRGSAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVOTA 1906
QY 1861 GSGLARIGLSWYADRPAGTYSGGNKRKLATALVGDPAVVPDEPTTGMPSARRFLWN 1920
DB 1907 GSGLARIGLSWYADRPAGTYSGGNKRKLATALVGDPAVVPDEPTTGMPSARRFLWN 1966
QY 1921 SLLAVREGSVMLTSHMEECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTVR 1980
DB 1967 SLLAVREGSVMLTSHMEECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTVR 2026
QY 1981 PAARSQPAAPAAVFAAEFFPGSELREAHGGRFLRQLPPGRCALARVFGELAVHGAEGVEDF 2040
DB 2027 PAARSQPAAPAAVFAAEFFPGSELREAHGGRFLRQLPPGRCALARVFGELAVHGAEGVEDF 2086
```

Qy 2041 SVSQTMLLEEVFLYFSKDGKDEDETEQKEAGVGVDPAPGLOHPKRVSQFLDDSTAETVL 2100
Db 2087 SVSQTMLLEEVFLYFSKDGKDEDETEQKEAGVGVDPAPGLOHPKRVSQFLDDSTAETVL 2146

RESULT 3
US-10-182-006-2
; Sequence 2, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10182,006
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: PCT/US01/02191
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-006-2

Query Match 99.98; Score 10893; DB 15; Length 2146;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2098; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPLEHHCHEPNKPLPSAGTVPWLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLSRLL 60
Db 47 PPLEHHCHEPNKPLPSAGTVPWLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLSRLL 106

Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMDVAELTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMDVAELTSLRT 166

Qy 121 ESIGLALGOQEPHLSLEAEADLAQELLALRSVELRALLORPGTSGPLELLEALCS 180
Db 167 ESIGLALGOQEPHLSLEAEADLAQELLALRSVELRALLORPGTSGPLELLEALCS 226

Qy 181 VRGSSSTVGPLNWEASDLMELVGQEPESALPDSSLSPACSELIIGALDGHPLSLRWRR 240
Db 227 VRGSSSTVGPLNWEASDLMELVGQEPESALPDSSLSPACSELIIGALDGHPLSLRWRR 286

Qy 241 LKPLILGKLLFAPDTPPTRKLMAQVNRFTBELTLRLDRVREWMELGPRIFTFMNDSNVA 300
Db 287 LKPLILGKLLFAPDTPPTRKLMAQVNRFTBELTLRLDRVREWMELGPRIFTFMNDSNVA 346

Qy 301 MLQRLLOQDEGRQPPGGRDHMEALRSFLDPGSGGYSQDAHADVGHVLTIGRVTEC 360
Db 347 MLQRLLOQDEGRQPPGGRDHMEALRSFLDPGSGGYSQDAHADVGHVLTIGRVTEC 406

Qy 361 LSLDLKLEAAPSEALVSRALQLLAHHPFWAGVWFLGPDSSDPTHEHTPDLGPGHVRKI 420
Db 407 LSLDLKLEAAPSEALVSRALQLLAHHPFWAGVWFLGPDSSDPTHEHTPDLGPGHVRKI 466

Qy 421 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 480
Db 467 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 526

Qy 481 LYLOQMPYPCYVDDVFLRVLRSRLPLFLTLAWIYSVTLTKVAVREKETRLRDTMRANGL 540
Db 527 LYLOQMPYPCYVDDVFLRVLRSRLPLFLTLAWIYSVTLTKVAVREKETRLRDTMRANGL 586

Qy 541 SRAVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPYSHPGVWFLFLAAFAVATVTSQFLL 600

Db 587 SRAVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPCSHPGVWFLFLAAFAVATVTSQFLL 646
Qy 601 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVWRDLRPLAGRVAASLLSPVAFGGCESLAL 660
Db 647 SAFFSRANLAAACGGLAYFSLYLYPYVLCVAVWRDLRPLAGRVAASLLSPVAFGGCESLAL 706

Qy 661 LEEQGECAQWHNVGTRPTADVFSIAQVSGILLDDAALYGLATWYLEAVCGQVGIPEPPWN 720
Db 707 LEEQGECAQWHNVGTRPTADVFSIAQVSGILLDDAALYGLATWYLEAVCGQVGIPEPPWN 766

Qy 721 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQPALRG 780
Db 767 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQPALRG 826

Qy 781 LSLDFYQGHITAFIGHNGAGKTTTILSGLPPPSGGSAPFILGHDVRSMAAIRPHLGVVC 840
Db 827 LSLDFYQGHITAFIGHNGAGKTTTILSGLPPPSGGSAPFILGHDVRSMAAIRPHLGVVC 886

Qy 841 POYNVLFDMLTVDHEVWFYGRGLGSAAVVGPQDRLLQDVGIVSKQSVOTRHLSGMQR 900
Db 887 POYNVLFDMLTVDHEVWFYGRGLGSAAVVGPQDRLLQDVGIVSKQSVOTRHLSGMQR 946

Qy 901 KLSVAIAFVGGSVWILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDAAELLD 960
Db 947 KLSVAIAFVGGSVWILDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDAAELLD 1006

Qy 961 RVAVAGGRLCCCGSPLFLRRHLSGYLTLVKARLPLTTNEKADTDMESVDTROKKN 1020
Db 1007 RVAVAGGRLCCCGSPLFLRRHLSGYLTLVKARLPLTTNEKADTDMESVDTROKKN 1066

Qy 1021 GSGSRVGTPLLALVQHWPVPGARLVEELPHELVLVLPYTGADHGGSPATIFRELDRLAE 1080
Db 1067 GSGSRVGTPLLALVQHWPVPGARLVEELPHELVLVLPYTGADHGGSPATIFRELDRLAE 1126

Qy 1081 LRLTGIGISDTSLEEIFLKVVEECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1140
Db 1127 LRLTGIGISDTSLEEIFLKVVEECAADTDMEDSCGQHLCTGIAGLDVTLRLKMPQETA 1186

Qy 1141 LENGEPAGSAPETDQSGDPDAVGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
Db 1187 LENGEPAGSAPETDQSGDPDAVGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1246

Qy 1201 ALFVGLALVFSLLVPPFGHPALRLSPTWYGAOVSFSEDPADPGDPRARLEALLQBAGL 1260
Db 1247 ALFVGLALVFSLLVPPFGHPALRLSPTWYGAOVSFSEDPADPGDPRARLEALLQBAGL 1306

Qy 1261 EEPVVOHSSHRFSAPEVAEVAKVLASGNWTPSPSPACQCSQFGARRLLPDCPAAAGGP 1320
Db 1307 EEPVVOHSSHRFSAPEVAEVAKVLASGNWTPSPSPACQCSQFGARRLLPDCPAAAGGP 1366

Qy 1321 PPQAVTGSGEVQNTLGRNLSDFLVKTYPRLVROGLTKTKWVNEVRYGGSFGRRDPGL 1380
Db 1367 PPQAVTGSGEVQNTLGRNLSDFLVKTYPRLVROGLTKTKWVNEVRYGGSFGRRDPGL 1426

Qy 1381 PSQOELGRSVEELWALLSPLPGCALDRVLKNLTAWAHSLEDAQDSLKLTFWNNKGHSWVAF 1440
Db 1427 PSQOELGRSVEELWALLSPLPGCALDRVLKNLTAWAHSLEDAQDSLKLTFWNNKGHSWVAF 1486

Qy 1441 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSICVVFAM 1500
Db 1487 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSICVVFAM 1546

Qy 1501 SFVPAFTVLVIBERVTRAKHLOIMGGLSPTLWLNFLNMDMKNLYLPACIVVLIFLAFQ 1560
Db 1547 SFVPAFTVLVIBERVTRAKHLOIMGGLSPTLWLNFLNMDMKNLYLPACIVVLIFLAFQ 1606

Qy 1561 QRAVAPANLPALELILLLGWSITPLMYPASFEFSPSTAYVVLTCINLFIGINGSMAT 1620
Db 1607 QRAVAPANLPALELILLLGWSITPLMYPASFEFSPSTAYVVLTCINLFIGINGSMAT 1666

Qy 1621 FVLELFSQDKLQBSRILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGRDQFQSPLR 1680
Db 1667 FVLELFSQDKLQBSRILKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGRDQFQSPLR 1726

QY 1681 WEVVGKLLAMVIOQPLPFLTLQHRSQLPQVRVSLPILGHEDEDVARERERVQ 1740
DB 1727 WEVVGKLLAMVIOQPLPFLTLQHRSQLPQVRVSLPILGHEDEDVARERERVQ 1786
QY 1741 ATQGDVLRNLTKVYRGQMPAVDRLCIGIPPGECFGLLGVNGAGKTSTFRMVTGDTLA 1800
DB 1787 ATQGDVLRNLTKVYRGQMPAVDRLCIGIPPGECFGLLGVNGAGKTSTFRMVTGDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
DB 1847 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1906
QY 1861 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARREFLWN 1920
DB 1907 GSGLARLGLSWADRPAGTYSGNKRKLATALLVGDPAVFLDPTTGMDSARREFLWN 1966
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
DB 1967 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQPAFAAFAEPGSELBEANGGRRLRFOLPGGRCALARVFGELAVHGAHGVDF 2040
DB 2027 PAARSQPAFAAFAEPGSELBEANGGRRLRFOLPGGRCALARVFGELAVHGAHGVDF 2086
QY 2041 SVSQTMLEEVFLYFSKQDKEDTEOEKAGVGVDPAPGLQHPKRVSOFLDDPSTAEIVL 2100
DB 2087 SVSQTMLEEVFLYFSKQDKEDTEOEKAGVGVDPAPGLQHPKRVSOFLDDPSTAEIVL 2146

RESULT 4

US-10-332-447-27
Sequence 27, Application US/10332447.
Publication No. US20040053258A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.; RAUMANN, Brigitte B.;
APPLICANT: THORNTON, Michael; DING, Li; YUE, Henry;
APPLICANT: TANG, Y. Tom; HARLAND, Lee; BURFORD, Neil;
APPLICANT: GREENE, Barrie D.; SANJANWALA, Madhu S.;
APPLICANT: BAUGHN, Marian R.; YAO, Monique G.; YANG, Junming;
APPLICANT: ARVIZU, Chandra S.; GANDHI, Ameena R.;
APPLICANT: HAPALIA, April J.A.; TRYBOULEY, Catherine M.;
APPLICANT: WALSH, Roderick T.; RAMKUMAR, Javalaxmi;
APPLICANT: LU, Yan; LU, Dying Aina M.; AZIMZAI, Yalda;
APPLICANT: LAJ, Preeti; SELLOTT, Vicki S.; NGUYEN, Damiel B.;
APPLICANT: KHAN, Farrah A.; KEARNEY, Liam; THANGAVELU, Kavitha;
APPLICANT: DAS, Debopriya; POLICKY, Jennifer L.
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
FILE REFERENCE: PI-0149 USN

CURRENT APPLICATION NUMBER: US/10/332,447

CURRENT FILING DATE: 2003-01-07

PRIOR APPLICATION NUMBER: US 60/216,547

PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: US 60/218,232

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/220,112

PRIOR FILING DATE: 2000-07-21

PRIOR APPLICATION NUMBER: US 60/221,839

PRIOR FILING DATE: 2000-07-28

NUMBER OF SEQ ID NOS: 64

SOFTWARE: PERL Program

SEQ ID NO 27

LENGTH: 2180

TYPE: PR

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20040053258A1 7475603CD1

US-10-332-447-27

Query Match

99.9%; Score 10883; DB 15; Length 2180;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFNKPLPSAGTVPWLQGLICNVNNTCFPQLTPGEEPRGRLSNFNDLSVRL 60
DB 81 PPLEHHECHFNKPLPSAGTVPWLQGLICNVNNTCFPQLTPGEEPRGRLSNFNDLSVRL 140
QY 61 ADARTVLGASAHRTLAGLGLKLIATLRAARSTAOQPQTQKSPLEPPMLDVAELLSTLRT 120
DB 141 ADARTVLGASAHRTLAGLGLKLIATLRAARSTAOQPQTQKSPLEPPMLDVAELLSTLRT 200
QY 121 ESLGIALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSGLLEALCS 180
DB 201 ESLGIALGQAQEPHLSLEAAEDLAQELLALRSVLVELRALLQRPRTGTSGLLEALCS 260
QY 181 VRGSPSTVGPSLNWYEASDLMELVQGBEPESALPDSSLPSPACSELIGALDHPHLSLLWRR 240
DB 261 VRGSPSTVGPSLNWYEASDLMELVQGBEPESALPDSSLPSPACSELIGALDHPHLSLLWRR 320
QY 241 LKPLTLGLKLLFAPDPTFTRKLMQAVNRTFEELTLRLDRVREVMELGPRIFTFMDSSNVA 300
DB 321 LKPLTLGLKLLFAPDPTFTRKLMQAVNRTFEELTLRLDRVREVMELGPRIFTFMDSSNVA 380
QY 301 MLQRLQLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 360
DB 381 MLQRLQLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 440
QY 361 LSLDKLEAAPSEALVSRALQLLAEHFWAGVVFGLGPEDDSDPTBHPDPLGPGHVRKI 420
DB 441 LSLDKLEAAPSEALVSRALQLLAEHFWAGVVFGLGPEDDSDPTBHPDPLGPGHVRKI 500
QY 421 RMDIDVTRTNKIRDRFWDGPAADPLTLRLRVWGGFVYLODLVERAAVRVLSGANPRAG 480
DB 501 RMDIDVTRTNKIRDRFWDGPAADPLTLRLRVWGGFVYLODLVERAAVRVLSGANPRAG 560
QY 481 LYLOQMPYPCYVDVFLVLSRSLPLFTLAWISVTLTVKAVVREKETRDLTRAMGL 540
DB 561 LYLOQMPYPCYVDVFLVLSRSLPLFTLAWISVTLTVKAVVREKETRDLTRAMGL 620
QY 541 SRAVLWIGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVVVFLPAAFAVATVQSFLL 600
DB 621 SRAVLWIGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVVVFLPAAFAVATVQSFLL 680
QY 601 SAFTSRANLAACGLAYFSLVPLVLCVNRDLRPGGRVAAASLLSPVAFPGCESLAL 660
DB 681 SAFTSRANLAACGLAYFSLVPLVLCVNRDLRPGGRVAAASLLSPVAFPGCESLAL 740
QY 661 LEEQEGQAQMNVTGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPWPN 720
DB 741 LEEQEGQAQMNVTGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGIPWPN 800
QY 721 PPFRRSYWCGPRPKSPAPCPTDLPKVLVEAPPGLSPPGVSVRSLEKRFPGSPQALRG 780
DB 801 PPFRRSYWCGPRPKSPAPCPTDLPKVLVEAPPGLSPPGVSVRSLEKRFPGSPQALRG 860
QY 781 LSLDPYQGHITAFIHNAGAGKTTTILSGLFPPSGGSATILGHVDRSSMAAIRHLGVC 840
DB 861 LSLDPYQGHITAFIHNAGAGKTTTILSGLFPPSGGSATILGHVDRSSMAAIRHLGVC 920
QY 841 PQYNVLFDMLTVDHVMFYGRKGLSAAVVGPEQDRLLQDVGLSVKQSVQTRHLSGGMR 900
DB 921 PQYNVLFDMLTVDHVMFYGRKGLSAAVVGPEQDRLLQDVGLSVKQSVQTRHLSGGMR 980
QY 901 KLSVAIAFPVGGSVQVILDEPTAGVDPPASRRGIWELLKLYREGRTIILSTHLLDEALLGD 960
DB 981 KLSVAIAFPVGGSVQVILDEPTAGVDPPASRRGIWELLKLYREGRTIILSTHLLDEALLGD 1040
QY 961 RVAVVAGRLCCGSPFLREHLSGYVLTAVKARLPLTTNEKADTDMEGSDVTRQKKN 1020
DB 1041 RVAVVAGRLCCGSPFLREHLSGYVLTAVKARLPLTTNEKADTDMEGSDVTRQKKN 1100
QY 1021 GSGQSRVGTGPOLLALVQHWVPGARLVBELPHELVLVLPYTGADHGSFATLRELDTRLAE 1080

Db 1101 GSQSRVGTQQLLALVOHWVPGARLVEELPHELVLVLPTGCAHDCGSPATLFRDLDTLAE 1160
Qy 1081 LRLTGYGISTSLBEEIFLKVVBECAADTMEDGSCGQHLCTGIAGLDVTLRLKMPPOETA 1140
Db 1161 LRLTGYGISTSLBEEIFLKVVBECAADTMEDGSCGQHLCTGIAGLDVTLRLKMPPOETA 1220
Qy 1141 LENGEPAAGSAPETDQSGSDPAVGRVQGWALTRQOLQALLKRFLLARRSRGLPAQIVLP 1200
Db 1221 LENGEPAAGSAPETDQSGSDPAVGRVQGWALTRQOLQALLKRFLLARRSRGLPAQIVLP 1280
Qy 1201 ALFVGLALVSLIYVPPGHPALRLSPTMYGAQVSFFSEADPGDGRARLLEALLQAGL 1260
Db 1281 ALFVGLALVSLIYVPPGHPALRLSPTMYGAQVSFFSEADPGDGRARLLEALLQAGL 1340
Qy 1261 EEPVQVHSSHRFSAPEVPAEYAKVLASGNMTPESSPACQCSQPGARRLLPDPCPAAAGGP 1320
Db 1341 EEPVQVHSSHRFSAPEVPAEYAKVLASGNMTPESSPACQCSQPGARRLLPDPCPAAAGGP 1400
Qy 1321 PPOQAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEVRYGGSFSLGGRDPGL 1380
Db 1401 PPOQAVTGSSEVVQNLTCRNLSDFLVKTYPRLVROGLKTKKWNVEVRYGGSFSLGGRDPGL 1460
Qy 1381 PSQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLOADSLKTFWNNKGWHSWAF 1440
Db 1461 PSQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLOADSLKTFWNNKGWHSWAF 1520
Qy 1441 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVFAM 1500
Db 1521 VNRSNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVFAM 1580
Qy 1501 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWLGHNFLDMCNLYVPACIVVLIFLAFQ 1560
Db 1581 SFVPASTVLVIERVTRAKHLQMGGLSPTLYWLGHNFLDMCNLYVPACIVVLIFLAFQ 1640
Qy 1561 QRAVVAANILPALLLLLLYGWSITPLMPYASFSSVPSTAYVVLTCINLFIGINGSMAT 1620
Db 1641 QRAVVAANILPALLLLLLYGWSITPLMPYASFSSVPSTAYVVLTCINLFIGINGSMAT 1700
Qy 1621 FVLELFSQDKLQEVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROFQSPLR 1680
Db 1701 FVLELFSQDKLQEVSRILKQVFLIFPHFCLGRGLIDMVNRQAMADAFERLGDROFQSPLR 1760
Qy 1681 WEVVGKLLAMVIOGPIFLFTALLQHSQLLPQPRVRSIPLIGEEDVARERERVQ 1740
Db 1761 WEVVGKLLAMVIOGPIFLFTALLQHSQLLPQPRVRSIPLIGEEDVARERERVQ 1820
Qy 1741 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCEFCGLLGVNGAGKTSFPMVGTDLA 1800
Db 1821 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGCEFCGLLGVNGAGKTSFPMVGTDLA 1880
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1860
Db 1881 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTA 1940
Qy 1861 GSGLARLGLSWYADRPAGTYSGGNKRKLATALVGPDPVAVVFLDEPTTGMDSARRFLWN 1920
Db 1941 GSGLARLGLSWYADRPAGTYSGGNKRKLATALVGPDPVAVVFLDEPTTGMDSARRFLWN 2000
Qy 1921 SLLAVVREGRSVMLTSHSMECEALCSRLATMNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 2001 SLLAVVREGRSVMLTSHSMECEALCSRLATMNGRFRCLGSPQHLKGRFAAGHTLTLRV 2060
Qy 1981 PAARSQAAAAFVAEEFPGSELREAHGGRRLRFQLPFGRCALARVFGELAVHGABSHGVEDF 2040
Db 2061 PAARSQAAAAFVAEEFPGSELREAHGGRRLRFQLPFGRCALARVFGELAVHGABSHGVEDF 2120
Qy 2041 SVSQTMLBEVFLYFSKQGDDETEEQKEAGVVDPAAGLOHPRKVSQFLLDDPSTAETVL 2100
Db 2121 SVSQTMLBEVFLYFSKQGDDETEEQKEAGVVDPAAGLOHPRKVSQFLLDDPSTAETVL 2180

; Sequence 2, Application US/09858194
; Patent No. US20020061590A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, MARIA
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF
; FILE REFERENCE: MNI-153
; CURRENT APPLICATION NUMBER: US/09/858,194
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/204,211
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-194-2

Query Match 99.9%; Score 10880; DB 9; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPLEHHECHFPNKPPLPSAGTVPMLOGLICNVNNTCPQLTPGEPGRLSNFNDSLVSRL 60
Db 45 PPLEHHECHFPNKPPLPSAGTVPMLOGLICNVNNTCPQLTPGEPGRLSNFNDSLVSRL 104
Qy 61 ADARTVLGGASAHRTLAGLKLATLRAARSTAAQOPTKQSPLEPMLDVAELLTSLRT 120
Db 105 ADARTVLGGASAHRTLAGLKLATLRAARSTAAQOPTKQSPLEPMLDVAELLTSLRT 164
Qy 121 ESGLGALQAOEPLHSLLEAAEDLAQELLALRSVELRALLQPRGTSGLLELSEALCS 180
Db 165 ESGLGALQAOEPLHSLLEAAEDLAQELLALRSVELRALLQPRGTSGLLELSEALCS 224
Qy 181 VRGSSSTVGPLSNWYEASDLMELVGOEPESALPDSSLSPACSELIGALDHPHLSRLWRR 240
Db 225 VRGSSSTVGPLSNWYEASDLMELVGOEPESALPDSSLSPACSELIGALDHPHLSRLWRR 284
Qy 241 LKPLILGLKLAFADPTPTTRKLMQAQVNTFEEELTLDRVREVMELGPRITTFMNDSSNVA 300
Db 285 LKPLILGLKLAFADPTPTTRKLMQAQVNTFEEELTLDRVREVMELGPRITTFMNDSSNVA 344
Qy 301 MLQRLLOMDEGRQPRPGGRDHMEALRSFLDPCSGGYSWQDAHADVGHVLTGRVTEC 360
Db 345 MLQRLLOMDEGRQPRPGGRDHMEALRSFLDPCSGGYSWQDAHADVGHVLTGRVTEC 404
Qy 361 LSLDKLEAAPSEALVSRLAQLLAHRFWAGVFLGPESSDPTHEPTDLPGLGPHVRIKI 420
Db 405 LSLDKLEAAPSEALVSRLAQLLAHRFWAGVFLGPESSDPTHEPTDLPGLGPHVRIKI 464
Qy 421 RMDIDVVTRTNKIDRPFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSCANPRAG 480
Db 465 RMDIDVVTRTNKIDRPFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSCANPRAG 524
Qy 481 LYLOQMPYPCYVDVDFLRLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDTRMAMGL 540
Db 525 LYLOQMPYPCYVDVDFLRLVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLDTRMAMGL 584
Qy 541 SRVILWLGWFLSCLGPPLLSAALLVLVLKLGDLIPYSHPGVVFLLAAFAVATVTSFLL 600
Db 585 SRVILWLGWFLSCLGPPLLSAALLVLVLKLGDLIPYSHPGVVFLLAAFAVATVTSFLL 644
Qy 601 SAFFSRANLAAACGLAYFSLYLPYVLCVAWRDLRPPAGRVAAASLLSPVAFGFCESIAL 660
Db 645 SAFFSRANLAAACGLAYFSLYLPYVLCVAWRDLRPPAGRVAAASLLSPVAFGFCESIAL 704
Qy 661 LEEQGEQAQHNVTGTRTADVFSLAQVSGLLLDALYGLATWYLEAVCGQVGI PBPWN 720
Db 705 LEEQGEQAQHNVTGTRTADVFSLAQVSGLLLDALYGLATWYLEAVCGQVGI PBPWN 764
Qy 721 FPFRRSYWCGRPPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 780

Db 765 FFRERSYMCPRPKSPAPCPPTLPDPKVLVEBAPPGLSPGVSVRSLEKRFPGSPQALRG 824
Qy 781 LSLDFVQGHITAPLGHGAGKTTLSLSGLPPSGGSFALIGHDVRSSMAIRPHLVC 840
Db 825 LSLDFVQGHITAPLGHGAGKTTLSLSGLPPSGGSFALIGHDVRSSMAIRPHLVC 884
Qy 841 PQYNVLFDMLTVDHFWFYGRLKGLSAAVVGPEQDRLLQDVGLSKVQSQVTRHLSGGMOR 900
Db 885 PQYNVLFDMLTVDHFWFYGRLKGLSAAVVGPEQDRLLQDVGLSKVQSQVTRHLSGGMOR 944
Qy 901 KLSVAIAFVGGQVILDEPTAGVDPASRRGIWELLKYREGRTILISTHLLDEALLGD 960
Db 945 KLSVAIAFVGGQVILDEPTAGVDPASRRGIWELLKYREGRTILISTHLLDEALLGD 1004
Qy 961 RVAVVAGRLCCGSPFLRRHLGSGYILTAKARLPLTNEKADTMEGSDVTROEKN 1020
Db 1005 RVAVVAGRLCCGSPFLRRHLGSGYILTAKARLPLTNEKADTMEGSDVTROEKN 1064
Qy 1021 GSGSRVGTQQLLALVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLFRDLRLAE 1080
Db 1065 GSGSRVGTQQLLALVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLFRDLRLAE 1124
Qy 1081 LRLTGYISDTSLEIFLKVBECAADTMDGSCQHLCGTGAGLDVTLRLKMPQETA 1140
Db 1125 LRLTGYISDTSLEIFLKVBECAADTMDGSCQHLCGTGAGLDVTLRLKMPQETA 1184
Qy 1141 LENGEPAGSAPETDGGSDPDAVGRVQGWALTRQQLQALLKBFLLARSRRLFAQIVLP 1200
Db 1185 LENGEPAGSAPETDGGSDPDAVGRVQGWALTRQQLQALLKBFLLARSRRLFAQIVLP 1244
Qy 1201 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVFFSEDPDGRARLLLEALLQEAGL 1260
Db 1245 ALFVGLALVFSILVPPFGHYPALRLSPMYGAQVFFSEDPDGRARLLLEALLQEAGL 1304
Qy 1261 BEPPVQSHSRPSADEVPAEVAKVASGNWTPESPSPACQSQPCQARLLPDCPAAAGGP 1320
Db 1305 BEPPVQSHSRPSADEVPAEVAKVASGNWTPESPSPACQSQPCQARLLPDCPAAAGGP 1364
Qy 1321 PPPQAVTSGGEVQNLTCRNLSDFLVKTYPRLVROGLTKKWNVEVRYGSGSLGRDPGL 1380
Db 1365 PPPQAVTSGGEVQNLTCRNLSDFLVKTYPRLVROGLTKKWNVEVRYGSGSLGRDPGL 1424
Qy 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSIDAQDSLKIFWNNKNGHSMVAF 1440
Db 1425 PSGQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSIDAQDSLKIFWNNKNGHSMVAF 1484
Qy 1441 VNRAASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVFPAM 1500
Db 1485 VNRAASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVFPAM 1544
Qy 1501 SFVPASFTVLVLEERVTRAKHLQLMGGLSPTLYWGNFLWMCNLYVPACIVVLIPLAFQ 1560
Db 1545 SFVPASFTVLVLEERVTRAKHLQLMGGLSPTLYWGNFLWMCNLYVPACIVVLIPLAFQ 1604
Qy 1561 QRAYVAPANLPALELLLYGWSITPLMYPASFFSPVSTAYVTLTINLFIGINGSNAT 1620
Db 1605 QRAYVAPANLPALELLLYGWSITPLMYPASFFSPVSTAYVTLTINLFIGINGSNAT 1664
Qy 1621 FVLELFSQKLQEVSRILKQVFLIPPHFLCGRGLIDMYRNQAMADAFRLGDRQFQSPLR 1680
Db 1665 FVLELFSQKLQEVSRILKQVFLIPPHFLCGRGLIDMYRNQAMADAFRLGDRQFQSPLR 1724
Qy 1681 WEVVGKLNLAAMVIOQPLFLFTLLQHRSQLLPQPRVRSPLPLGGEDEVAREERVQOG 1740
Db 1725 WEVVGKLNLAAMVIOQPLFLFTLLQHRSQLLPQPRVRSPLPLGGEDEVAREERVQOG 1784
Qy 1741 ATQGDVLVRLNLTKYRQORMPAVDRLCLGIPPGECFGLLVNGAGKTSITFRMTVGTDLA 1800
Db 1785 ATQGDVLVRLNLTKYRQORMPAVDRLCLGIPPGECFGLLVNGAGKTSITFRMTVGTDLA 1844
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLURGYPEAQVQTA 1860
Db 1845 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLURGYPEAQVQTA 1904

Qy 1861 GSGLARLGLSMYADRPACTYSGNKRKLATALLVGDPAVVPFLDEPTTGMDSABRRFLWN 1920
Db 1905 GSGLARLGLSMYADRPACTYSGNKRKLATALLVGDPAVVPFLDEPTTGMDSABRRFLWN 1964
Qy 1921 SILAVVREGRSVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1965 SILAVVREGRSVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2024
Qy 1981 PAARSQAAAAFVAABFPFPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAHGVDF 2040
Db 2025 PAARSQAAAAFVAABFPFPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAHGVDF 2084
Qy 2041 SVSQTMLLEEVFLYFSKOQKDEDETEOEKAGVGVDPAFCLOHPKEVSOFLDDPSTAEVTL 2100
Db 2085 SVSQTMLLEEVFLYFSKOQKDEDETEOEKAGVGVDPAFCLOHPKEVSOFLDDPSTAEVTL 2144

RESULT 6
US-10-154-419-2
; Sequence 2, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8059, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FHS58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-2

Query Match 99.9%; Score 10880; DB 14; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PPLEHHECHFPNKPPLPSAGTVPMLOGLICNVNNTCFPQLTGEEFGRLSNFNDSLVSRLL 60
Db 45 PPLEHHECHFPNKPPLPSAGTVPMLOGLICNVNNTCFPQLTGEEFGRLSNFNDSLVSRLL 104
Qy 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPOPTKOSPPLPPMLDVAELLTSLRT 120
Db 105 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPOPTKOSPPLPPMLDVAELLTSLRT 164
Qy 121 ESLGLALCQAEPLHSLLEAAEDLAQELLARLSVELRALLQRPRTSGPRLLEALCS 180
Db 165 ESLGLALCQAEPLHSLLEAAEDLAQELLARLSVELRALLQRPRTSGPRLLEALCS 224
Qy 181 VRGSPSTVPSLNWYEASDLMELVQBPESALPOSSISLSPACSELIGALDHPSLRLLWR 240
Db 225 VRGSPSTVPSLNWYEASDLMELVQBPESALPOSSISLSPACSELIGALDHPSLRLLWR 284
Qy 241 LKPLILGKLFPAPDTPTRKLMQAVNRTFEELTLRDVREVMELGPRIFTFMDSSNVA 300
Db 285 LKPLILGKLFPAPDTPTRKLMQAVNRTFEELTLRDVREVMELGPRIFTFMDSSNVA 344
Qy 301 MLQRLQWQDEGRQRPGRGRDHMEALRSFLDPGSGGYSMQADAHADVGLVGLTGRVTEC 360
Db 345 MLQRLQWQDEGRQRPGRGRDHMEALRSFLDPGSGGYSMQADAHADVGLVGLTGRVTEC 404
Qy 361 LSLDKLEAAPSEAAALVSRALQLLAEHRFWAGVFLGPEDSDPTTEHPTDLPQGHVRIKI 420
Db 405 LSLDKLEAAPSEAAALVSRALQLLAEHRFWAGVFLGPEDSDPTTEHPTDLPQGHVRIKI 464


```
Qy 421 RMDIDVTRTKIRDRFWDPGPADPLTDLRYYWGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 465 RMDIDVTRTKIRDRFWDPGPADPLTDLRYYWGGFVYLQDLVERAAVRVLSGANPRAG 524
Qy 481 LYLOMPYPCVDDVFURVSRPLFLTLAWIYSVTLTKAVVREKETRLDTRAMGL 540
Db 525 LYLOMPYPCVDDVFURVSRPLFLTLAWIYSVTLTKAVVREKETRLDTRAMGL 584
Qy 541 SRAVLWLGWFLSLGPFLLSALLVLVLKGLDILPYSHPGVVFLLAFAFVATVTSFLL 600
Db 585 SRAVLWLGWFLSLGPFLLSALLVLVLKGLDILPYSHPGVVFLLAFAFVATVTSFLL 644
Qy 601 SAFPSTRANLAAACGLAYFSLYLPLYLVCVWRDLRPPAGGRVAASLLSPVAFGFCESLAL 660
Db 645 SAFPSTRANLAAACGLAYFSLYLPLYLVCVWRDLRPPAGGRVAASLLSPVAFGFCESLAL 704
Qy 661 LEEQEGQAHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPNW 720
Db 705 LEEQEGQAHNVGTRPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPNW 764
Qy 721 FPFRRSYWCGPRPKSPAPCTPLDPKVLVEEAPPGLSPGVSRSEKRPFGSPQPALRG 780
Db 765 FPFRRSYWCGPRPKSPAPCTPLDPKVLVEEAPPGLSPGVSRSEKRPFGSPQPALRG 824
Qy 781 LSLDFYQGHITAFIAGHNGAGTKTTLSILSGLFPSPGSGSAFTLGHDRSSMAAIRPHIGVC 840
Db 825 LSLDFYQGHITAFIAGHNGAGTKTTLSILSGLFPSPGSGSAFTLGHDRSSMAAIRPHIGVC 884
Qy 841 PQYNVLFDMLTVDHNVFYGRLKGLSAAVCGPEODRLLODGLVSKOSVOTRHLSSGMQR 900
Db 885 PQYNVLFDMLTVDHNVFYGRLKGLSAAVCGPEODRLLODGLVSKOSVOTRHLSSGMQR 944
Qy 901 KLSVAIAFVGGSQVVLDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDAEALLGD 960
Db 945 KLSVAIAFVGGSQVVLDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDAEALLGD 1004
Qy 961 RVAVVAGRLCCCGSPILFLRHLSGGYTLTVKARLPLTNEKADTMEGSDVTRQSKKN 1020
Db 1005 RVAVVAGRLCCCGSPILFLRHLSGGYTLTVKARLPLTNEKADTMEGSDVTRQSKKN 1064
Qy 1021 GSQSRVGTGTPOLLALVOHWPGARLVEELPHELVLVLPYTGADHSGSATLIFRELDTRLAE 1080
Db 1065 GSQSRVGTGTPOLLALVOHWPGARLVEELPHELVLVLPYTGADHSGSATLIFRELDTRLAE 1124
Qy 1081 LRLTGYGIDSLSBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLLRLKMPQETA 1140
Db 1125 LRLTGYGIDSLSBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLLRLKMPQETA 1184
Qy 1141 LENGEPAAGSAPETDQGGSPDAVGRVQGWALTRQOLQALLKRLFLARRSRRLGPAQIVLP 1200
Db 1185 LENGEPAAGSAPETDQGGSPDAVGRVQGWALTRQOLQALLKRLFLARRSRRLGPAQIVLP 1244
Qy 1201 ALFVGLALVFSLIYVPPGHYPALRLSPMTYGAQVSPFSEDAPGPGGARLLLEALLQAGL 1260
Db 1245 ALFVGLALVFSLIYVPPGHYPALRLSPMTYGAQVSPFSEDAPGPGGARLLLEALLQAGL 1304
Qy 1261 EEPVQVHSSHRFSAPEVPAEYAKVLASGNMTTPESPACQCSQPGARRLLPDCPAAAGGP 1320
Db 1305 EEPVQVHSSHRFSAPEVPAEYAKVLASGNMTTPESPACQCSQPGARRLLPDCPAAAGGP 1364
Qy 1321 PPPQAVTGSBVGWQNLGTRNLSDFLVITYPRLVQGLTKKXWNEVRYGGFSLGGRDPGL 1380
Db 1365 PPPQAVTGSBVGWQNLGTRNLSDFLVITYPRLVQGLTKKXWNEVRYGGFSLGGRDPGL 1424
Qy 1381 PSGQELGRSVEELWALLSPILPGGALDRVLKNTAWAHSDDAQDSLKIWFNKGHWSVAF 1440
Db 1425 PSGQELGRSVEELWALLSPILPGGALDRVLKNTAWAHSDDAQDSLKIWFNKGHWSVAF 1484
Qy 1441 VNRSASNAIRLAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVVFAM 1500
Db 1485 VNRSASNAIRLAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVICVVFAM 1544
```

```
Qy 1501 SFVPASFTLVLIBERVTRAKHLOLMGGLSPTLYMLGNFLWDMCNLYLPACIVLVILFAFQ 1560
Db 1545 SFVPASFTLVLIBERVTRAKHLOLMGGLSPTLYMLGNFLWDMCNLYLPACIVLVILFAFQ 1604
Qy 1561 QRAYVAPANUPALLLLLLLXGWSITPLMYPASFFSPSTAYVVLTCINLIFIGINGSMAT 1620
Db 1605 QRAYVAPANUPALLLLLLLXGWSITPLMYPASFFSPSTAYVVLTCINLIFIGINGSMAT 1664
Qy 1621 FVLEFSDQKLQSVRILKQVFLIFPHFCIGRGLIDMVRNQAMADAFERLGDROFQSPLR 1680
Db 1665 FVLEFSDQKLQSVRILKQVFLIFPHFCIGRGLIDMVRNQAMADAFERLGDROFQSPLR 1724
Qy 1681 WEVVGKNLLAMVIOGPFLLFTLLQHRSQLLPQPRVRSIPLIGEEDEYVARERERVQ 1740
Db 1725 WEVVGKNLLAMVIOGPFLLFTLLQHRSQLLPQPRVRSIPLIGEEDEYVARERERVQ 1784
Qy 1741 ATQGDVLVLRNLTKVYRGQRMPPAVDRCLGIPGCECFGLLVNGAGTKSTFRMVGTDLA 1800
Db 1785 ATQGDVLVLRNLTKVYRGQRMPPAVDRCLGIPGCECFGLLVNGAGTKSTFRMVGTDLA 1844
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTA 1860
Db 1845 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTA 1904
Qy 1861 GSGLARLGLSWYADRPAAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDDPSARRFLWN 1920
Db 1905 GSGLARLGLSWYADRPAAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDDPSARRFLWN 1964
Qy 1921 SLIAVVRGSRVMLTSHSMECEALCSRLAIWNGRPRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1965 SLIAVVRGSRVMLTSHSMECEALCSRLAIWNGRPRCLGSPQHLKGRFAAGHTLTLRV 2024
Qy 1981 PAARSQPAAPAAFAAEPPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHBGVDEF 2040
Db 2025 PAARSQPAAPAAFAAEPPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAHBGVDEF 2084
Qy 2041 SVSQTMLVEEFLYFSKQDQKDETEOKEAGVGDPAQGLQHPKRVSQFLDDPSTAEVTVL 2100
Db 2085 SVSQTMLVEEFLYFSKQDQKDETEOKEAGVGDPAQGLQHPKRVSQFLDDPSTAEVTVL 2144
```

RESULT 7
US-10-618-281-44
; Sequence 44, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-44

Query Match 99.8%; Score 10874.5; DB 17; Length 2147;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```
Qy 1 PPLEHHECHFPNKPLPSAGTVPMVQLGLICNVNNTCFPQLTPGEBPGRLSFNDSLSVRL 60
Db 47 PPLEHHECHFPNKPLPSAGTVPMVQLGLICNVNNTCFPQLTPGEBPGRLSFNDSLSVRL 106
```

QY 61 ADARTVLGASAHRTLAGKLIATLRAARSTAOQPTKQSPLEPMMLDVAELLTSLRT 120
Db 107 ADARTVLGASAHRTLAGKLIATLRAARSTAOQPTKQSPLEPMMLDVAELLTSLRT 166
QY 121 ESLGALGQAQEPHLSLEAAEDIAQELLALRSVLRAALLORPGTSGPLELSEALCS 180
Db 167 ESLGALGQAQEPHLSLEAAEDIAQELLALRSVLRAALLORPGTSGPLELSEALCS 226
QY 181 VRGSPSTVGPNSLWYEAEDLMELVQEPESALPDSSLSPPACSELIAGDLSHPLSLLWRR 240
Db 227 VRGSPSTVGPNSLWYEAEDLMELVQEPESALPDSSLSPPACSELIAGDLSHPLSLLWRR 286
QY 241 LKPLILGKLLFAPDPTFFTRKMAQVNRTFEELTLRDVREWEMLGPRIFTFMDNSSVA 300
Db 287 LKPLILGKLLFAPDPTFFTRKMAQVNRTFEELTLRDVREWEMLGPRIFTFMDNSSVA 346
QY 301 MLORLLQWQDEGRROPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTEC 360
Db 347 MLORLLQWQDEGRROPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTEC 406
QY 361 LSLDKLEAAPSEAAALVSALQLLAHRFWAGVVFGLGEDSSDPTBHPPTDLGPGHVRKI 420
Db 407 LSLDKLEAAPSEAAALVSALQLLAHRFWAGVVFGLGEDSSDPTBHPPTDLGPGHVRKI 466
QY 421 RMDIDVTRTKIRDRFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 467 RMDIDVTRTKIRDRFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPYPCYVDDVFLRSLPLFLTLAMIYSVTITVKAIVREKETRIIRDNRAMGL 540
Db 527 LYLOQMPYPCYVDDVFLRSLPLFLTLAMIYSVTITVKAIVREKETRIIRDNRAMGL 586
QY 541 SRAVLWLGWFLSCLGPPFLSAAALVLKLGDIILPYSHPGVVFLFLAFAVATVTSPELL 600
Db 587 SRAVLWLGWFLSCLGPPFLSAAALVLKLGDIILPYSHPGVVFLFLAFAVATVTSPELL 646
QY 601 SAFTSRANLAAACGLAYFSLPYLVLCVAMRDLRPPAGGRVAASLLSPVAFGFCESIAL 660
Db 647 SAFTSRANLAAACGLAYFSLPYLVLCVAMRDLRPPAGGRVAASLLSPVAFGFCESIAL 706
QY 661 LEEQEGQWNVTRPTADVFLSAQUSGLILLDDAALYGLATWYLEAVCPQOYGIPEPWN 720
Db 707 LEEQEGQWNVTRPTADVFLSAQUSGLILLDDAALYGLATWYLEAVCPQOYGIPEPWN 766
QY 721 PPFRRSYWCGPRPKSPAPCTPDLDPKVLVEAPPLSGVSVRSLEKRFPGSPQALRG 780
Db 767 PPFRRSYWCGPRPKSPAPCTPDLDPKVLVEAPPLSGVSVRSLEKRFPGSPQALRG 826
QY 781 LSLDFYQGHITAFLGHNGAGKTTTILSILSGLFPFPGGSFAFILGHIDVRSMAAIRPHLVC 840
Db 827 LSLDFYQGHITAFLGHNGAGKTTTILSILSGLFPFPGGSFAFILGHIDVRSMAAIRPHLVC 886
QY 841 PQYNVLPMLTVDHEWVYGBLKGLSAAVPEQDRLLQDVLVSQSVQTRHLSGGMQR 900
Db 887 PQYNVLPMLTVDHEWVYGBLKGLSAAVPEQDRLLQDVLVSQSVQTRHLSGGMQR 946
QY 901 KLSVAIAFVGGSQVVLDEPTAGVDPPASRRGIWELLKLYREGRTILSTHLDRAELLGD 960
Db 947 KLSVAIAFVGGSQVVLDEPTAGVDPPASRRGIWELLKLYREGRTILSTHLDRAELLGD 1006
QY 961 RVAVAGRLCCCGSPLFLRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTRQEKKN 1020
Db 1007 RVAVAGRLCCCGSPLFLRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTRQEKKN 1066
QY 1021 GSQSRVGTPOQLLALVQHWPFGARLVEBELPHELVLVLPYTGANDGSPATLFRDLTRLAE 1080
Db 1067 GSQSRVGTPOQLLALVQHWPFGARLVEBELPHELVLVLPYTGANDGSPATLFRDLTRLAE 1126
QY 1081 LRLTYGIGSDTSLBEEIPLKVVVECAADTMDGSCGQHLCTGIGAGLDVTLRLKMPPOETA 1140
Db 1127 LRLTYGIGSDTSLBEEIPLKVVVECAADTMDGSCGQHLCTGIGAGLDVTLRLKMPPOETA 1186
QY 1141 LENGEPAGSAPETDQSGPDVAVRVOGWALTRQQLQALLKRLFLARRSRRLGFLAQIVLP 1200

Db 1187 LENGEPAGSAPETDQSGPDVAVRVOGWALTRQQLQALLKRLFLARRSRRLGFLAQIVLP 1246
QY 1201 ALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVSPFSEDAPGDPGRARLLLEALLQEAGL 1260
Db 1247 ALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVSPFSEDAPGDPGRARLLLEALLQEAGL 1306
QY 1261 BEPVPQSHSSHRFSAPEVPAEAKVLASGNWTPPEPSPACQSQOPCARLLPDCPAAAGGP 1320
Db 1307 BEPVPQSHSSHRFSAPEVPAEAKVLASGNWTPPEPSPACQSQOPCARLLPDCPAAAGGP 1366
QY 1321 PPPQAVTSGSEVWQNLTRGNLSDFLVKTYPRLVQGLTKKWKVNEVRYGGFSLGGDRDGL 1380
Db 1367 PPPQAVTSGSEVWQNLTRGNLSDFLVKTYPRLVQGLTKKWKVNEVRYGGFSLGGDRDGL 1426
QY 1381 PSGQELGRSVEBELWALLSPLPGGALDRVLKNTAWAHSLDAQDSLKIFNFKNGHSMWAF 1440
Db 1427 PSGQELGRSVEBELWALLSPLPGGALDRVLKNTAWAHSLDAQDSLKIFNFKNGHSMWAF 1486
QY 1441 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAM 1500
Db 1487 VNRSNATLRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAM 1546
QY 1501 SFVPASPTLVLIEBVRTAKHLQMGGLSPTLYWLGNEFLWDMCNLYLPACIVVLIFLAFQ 1560
Db 1547 SFVPASPTLVLIEBVRTAKHLQMGGLSPTLYWLGNEFLWDMCNLYLPACIVVLIFLAFQ 1606
QY 1561 QRAYVAPANLPAALLLILLYGWSITPLMYPASFPFSPSTAYVVLVTCINLFIGINGSWAT 1620
Db 1607 QRAYVAPANLPAALLLILLYGWSITPLMYPASFPFSPSTAYVVLVTCINLFIGINGSWAT 1666
QY 1621 FVLELFSQ-OQLQEVSRILKOVFLIFPHFCLGRGLIDMVRNOAMADAPERLGDROFQSP 1679
Db 1667 FVLELFSQ-OQLQEVSRILKOVFLIFPHFCLGRGLIDMVRNOAMADAPERLGDROFQSP 1726
QY 1680 RWEVVGKLLAMVIOGPLFTLLQLHRSQLLPQPRVRSPLLGEDEDEVARERERVVQ 1739
Db 1727 RWEVVGKLLAMVIOGPLFTLLQLHRSQLLPQPRVRSPLLGEDEDEVARERERVVQ 1786
QY 1740 GATQGDVLVRLNLTKVYRGQMPADVRLCLGIPGCEGGLGVNGAGTKSTFRMWTDGTL 1799
Db 1787 GATQGDVLVRLNLTKVYRGQMPADVRLCLGIPGCEGGLGVNGAGTKSTFRMWTDGTL 1846
QY 1800 ASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLRGVPEAQVAQT 1859
Db 1847 ASRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLRGVPEAQVAQT 1906
QY 1860 AGSGLARLGLSWYADRPAQTYSGGNKRLATALALVGDPAVYVFLDEPTTGMDDPSARRFLW 1919
Db 1907 AGSGLARLGLSWYADRPAQTYSGGNKRLATALALVGDPAVYVFLDEPTTGMDDPSARRFLW 1966
QY 1920 NSLLAVVREGSVMLTSHSMBECEALCSRLAIMNGRPRCLGSPQHLKGRFAAGHTLTLR 1979
Db 1967 NSLLAVVREGSVMLTSHSMBECEALCSRLAIMNGRPRCLGSPQHLKGRFAAGHTLTLR 2026
QY 1980 VPAARSPAAAFAAEPPGSELREAHGRLRFQPLPGGRCALARVFGELAVHGAHGVED 2039
Db 2027 VPAARSPAAAFAAEPPGSELREAHGRLRFQPLPGGRCALARVFGELAVHGAHGVED 2086
QY 2040 FSVSQTMLVEEVLYFSKXQDKDEDETEBQKEAGVGDVPAPGLQHPKRVSQFLLDDPSTASTV 2099
Db 2087 FSVSQTMLVEEVLYFSKXQDKDEDETEBQKEAGVGDVPAPGLQHPKRVSQFLLDDPSTASTV 2146
QY 2100 L 2100
Db 2147 L 2147

APPLICANT: Guo, Xiaojia
APPLICANT: Kekuda, Rameesh
APPLICANT: Miller, Charles E.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Patturajan, Meera
APPLICANT: Liu, Zhaocong
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Li, Li
APPLICANT: Vernet, Corine
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glendna
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shimkets, Richard A.
APPLICANT: Gangolli, Esha A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Caeman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liete, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
TITLE OF INVENTION: No. US20040030110alel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-322C
CURRENT APPLICATION NUMBER: US/10/114.270
PRIOR FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 470
SEQ ID NO 176
LENGTH: 2059
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-176

Query Match 95.4%; Score 10392.5; DB 15; Length 2059;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 PPLEHHECHPNKPLPSAGTVPWVLOGLICNVNNTCFPQLTPGEPGRLSNFNDLSVRL 60
Db 47 PPLEHHECHPNKPLPSAGTVPWVLOGLICNVNNTCFPQLTPGEPGRLSNFNDLSVRL 106
Qy 61 ADARTVLGGASAHRTLAGLKIATLAARSTAOPTKOSPLPPMLDVALLTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLKIATLAARSTAOPTKOSPLPPMLDVALLTSLRT 166

Qy 121 ESGLGLAQAEPLHSLLEAAEDLAQELALRSLVELRALLQPRGTSGPGLLELSEALCS 180
Db 167 ESGLGLAQAEPLHSLLEAAEDLAQELALRSLVELRALLQPRGTSGPGLLELSEALCS 226
Qy 181 VRGPGSTVGPSTLNWYEASDLMELVQEPESALPDSSISPAACSELIGALDGHPSRLLR 240
Db 227 VRGPGSTVGPSTLNWYEASDLMELVQEPESALPDSSISPAACSELIGALDGHPSRLLR 286
Qy 241 LKPLILGKLLFAPDTPFTRKMAQVNRTPFELTLLRDVREVWEMLGPRIETFMNDSSNVA 300
Db 287 LKPLILGKLLFAPDTPFTRKMAQVNRTPFELTLLRDVREVWEMLGPRIETFMNDSSNVA 346
Qy 301 MLQRLQMDQEGRRQPRPGGRDHEALRSLDPGSGYSWQDAHADVGHVLGTRVTEC 360
Db 347 MLQRLQMDQEGRRQPRPGGRDHEALRSLDPGSGYSWQDAHADVGHVLGTRVTEC 406
Qy 361 LSLDKLEAPSEALYSRALQLLAHRFWAGVVFPLGPEDSSDTEHPTDPLGCHVRIKI 420
Db 407 LSLDKLEAPSEALYSRALQLLAHRFWAGVVFPLGPEDSSDTEHPTDPLGCHVRIKI 466
Qy 421 RMDIDVVRTNKIRDFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAG 480
Db 467 RMDIDVVRTNKIRDFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRLSGANPRAG 526
Qy 481 LYLQMPYPCYVDDVFLRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRAMGL 540
Db 527 LYLQMPYPCYVDDVFLRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRAMGL 586
Qy 541 SRVLMLGFWFLSCIGPFLSAALLVLKLGDTLPYSHGCVFLFLAFAVAVTQSFLL 600
Db 587 SRVLMLGFWFLSCIGPFLSAALLVLKLGDTLPYSHGCVFLFLAFAVAVTQSFLL 646
Qy 601 SAFFSRANLAAACGLAYFSLYLPYVLCVAVRDLRPAAGRAVAAASLLSPVAFGCESSAL 660
Db 647 SAFFSRANLAAACGLAYFSLYLPYVLCVAVRDLRPAAGRAVAAASLLSPVAFGCESSAL 706
Qy 661 LEEGEGAGQHNVTGRTADVFLSAQVSGLLLLDAALYGLATWYLEAVCGQYGIPEPNW 720
Db 707 LEEGEGAGQHNVTGRTADVFLSAQVSGLLLLDAALYGLATWYLEAVCGQYGIPEPNW 766
Qy 721 FPERRSYWCGRPKSPAPCPTPLDPKVLVEEAPPGLSGVSVRSLEKRPFGSPQPALRG 780
Db 767 FPERRSYWCGRPKSPAPCPTPLDPKVLVEEAPPGLSGVSVRSLEKRPFGSPQPALRG 826
Qy 781 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGSAFILGHDVRSMAAIRPHLGV 840
Db 827 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGSAFILGHDVRSMAAIRPHLGV 886
Qy 841 PQYNVLFDMLTVDHEVWFYGRLLKGLSAAVVGPPQDRLLQVGLVSKOSVQTRHLSGMQR 900
Db 887 PQYNVLFDMLTVDHEVWFYGRLLKGLSAAVVGPPQDRLLQVGLVSKOSVQTRHLSGMQR 946
Qy 901 KLSVATAFVGGSOVITLDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDAAELLGD 960
Db 947 KLSVATAFVGGSOVITLDEPTAGVDPASRRGIWELLKYREGRTLILSTHLDAAELLGD 1006
Qy 961 RVAVVAGGRLLCCCGSPFLFRRHLSGGYLLTVKARLPLTTNEKADTMEGSDVTRQKKN 1020
Db 1007 RVAVVAGGRLLCCCGSPFLFRRHLSGGYLLTVKARLPLTTNEKADTMEGSDVTRQKKN 1066
Qy 1021 GSGQSRVGPQLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLDTLAE 1080
Db 1067 GSGQSRVGPQLLALVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLDTLAE 1126
Qy 1081 LRLTGIGISDTSLEEIFLKVVECAADTMEGSCGQHLCTAGLDVTLRLKMPPOETA 1140
Db 1127 LRLTGIGISDTSLEEIFLKVVECAADTMEGSCGQHLCTAGLDVTLRLKMPPOETA 1186
Qy 1141 LENGEPAAGSAPETDQSGSDPAVGRVQGWALTRQOLQALLKRLFLARRSRGLFAQIVLP 1200
Db 1187 LENGEPAAGSAPETDQSGSDPAVGRVQGWALTRQOLQALLKRLFLARRSRGLFAQIVLP 1246

QY 1201 ALFVGLALVFSILVPPFGHYPALRLSPMTMYGAQVSFFSDEADPGDPCGRARLLLEALLQEAGL 1260
DB 1247 ALFVGLALVFSILVPPFGHYPALRLSPMTMYGAQVSFFSDEADPGDPCGRARLLLEALLQEAGL 1306
QY 1261 BEPPVQHSRHSAPAEVPAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1320
DB 1307 BEPPVQHSRHSAPAEVPAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1366
QY 1321 PPQAVTSGEVVQNLTRGNLSDFLVKTYPRLVROGLTKKWNVRYGGFSLGGRDPGL 1380
DB 1367 PPQAVTSGEVVQNLTRGNLSDFLVKTYPRLVROGLTKKWNVRYGGFSLGGRDPGL 1426
QY 1381 PSQELGRSVEELWALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIWFNKNKGWHSMAVF 1440
DB 1427 PSQELGRSVEELWALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIWFNKNKGWHSMAVF 1486
QY 1441 VNRAENAILRAHLPFGPARHAHSITLNLHPLNLTKQELSEAAALMASSVDVLVSIICVVFAM 1500
DB 1487 VNRAENAILRAHLPFGPARHAHSITLNLHPLNLTKQELSEAAALMASSVDVLVSIICVVFAM 1546
QY 1501 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLNFLWDMCNVLPACIVULIFLAPQ 1560
DB 1547 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLNFLWDM----- 1588
QY 1561 QRAYVAPANLPALLLLLLLYGWSITPLMYPASFFSFVSTAYVVLTCINLFIGINGSMAT 1620
DB 1589 ----- 1588
QY 1621 FVLELFSQKLOEVSRILKQVFLIPPHFCCLGRGLIDMVRNQAMADAFERLGDROQSPPLR 1680
DB 1589 -----KLQEVSRILKQVFLIPPHFCCLGRGLIDMVRNQAMADAFERLGDROQSPPLR 1639
QY 1681 WEVVGKLLAMVIOQPLFLLTLLQHSQLLPOPRVSLPLGDEEDVAREERVVOG 1740
DB 1640 WEVVGKLLAMVIOQPLFLLTLLQHSQLLPOPRVSLPLGDEEDVAREERVVOG 1699
QY 1741 ATQGDVLVRLNLTKYRQORPAVDRLCLGIPPGCEGGLGVNGAGKSTTFRMVGTDLA 1800
DB 1700 ATQGDVLVRLNLTKYRQORPAVDRLCLGIPPGCEGGLGVNGAGKSTTFRMVGTDLA 1759
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
DB 1760 SRGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVQTA 1819
QY 1861 GSGLARLGLSWYADRPAGTYSGGNKRLATALNGDPAVVFLDPEPTTGMDPSARRFLWN 1920
DB 1820 GSGLARLGLSWYADRPAGTYSGGNKRLATALNGDPAVVFLDPEPTTGMDPSARRFLWN 1879
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
DB 1880 SLLAVVREGSRVMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1939
QY 1981 PAARSQPAFAAFAVFAEPFSGELREAHGGRRLRQLPPGRCALARVFGELAVHGAHGVEDF 2040
DB 1940 PAARSQPAFAAFAVFAEPFSGELREAHGGRRLRQLPPGRCALARVFGELAVHGAHGVEDF 1999
QY 2041 SVSQTMLEEVFLYFSKQDKEDTEEQEAGVGDVPAPGLQHPKRVSFQFLDDPSTAETVL 2100
DB 2000 SVSQTMLEEVFLYFSKQDKEDTEEQEAGVGDVPAPGLQHPKRVSFQFLDDPSTAETVL 2059

RESULT 9

US-10-182-006-4
; Sequence 4, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED THEREFROM
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: PCT/US01/02191

; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1873
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-006-4

Query Match 84.6%; Score 9213.5; DB 15; Length 1873;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1789; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 PPLEHHECHFPNKPSPAGTVPMLQGLICNVNNTCFPOLTEGEEFGRSLNFNDSIVSRLL 60
DB 47 PPLEHHECHFPNKPSPAGTVPMLQGLICNVNNTCFPOLTEGEEFGRSLNFNDSIVSRLL 106
QY 61 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLIRT 120
DB 107 ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLIRT 166
QY 121 ESLGLALQAOEPLHSLLEAAEDLAQELLALRSIVELRALLQRPGTSGPLELSEALCS 180
DB 167 ESLGLALQAOEPLHSLLEAAEDLAQELLALRSIVELRALLQRPGTSGPLELSEALCS 226
QY 181 VRGSSVTGSPSLNWEASDLMEYQGBEPESALPDSSLSPPACSELTIGALDHSPLSLLWRR 240
DB 227 VRGSSVTGSPSLNWEASDLMEYQGBEPESALPDSSLSPPACSELTIGALDHSPLSLLWRR 286
QY 241 LKPLILGKLLPAPDPTPTTRKLMQVNRTEBELTLRLDRVREVMELGPRIFTFMDNDSNVA 300
DB 287 LKPLILGKLLPAPDPTPTTRKLMQVNRTEBELTLRLDRVREVMELGPRIFTFMDNDSNVA 346
QY 301 MLQRLLOMDQRRQPRPGGRDHMEALRSFLDPCGGYSWODAHADVGLVGTIGRVTEC 360
DB 347 MLQRLLOMDQRRQPRPGGRDHMEALRSFLDPCGGYSWODAHADVGLVGTIGRVTEC 406
QY 361 LSLDKLEAAPSEAAALVSRAQLLAEHRFWAGVFLGPEDDSDPTHEPTDLPDGPCHVRIKI 420
DB 407 LSLDKLEAAPSEAAALVSRAQLLAEHRFWAGVFLGPEDDSDPTHEPTDLPDGPCHVRIKI 466
QY 421 RMDIDVTRTNKIRDRFWDPGPAADPLTDLYRWGGGFYLLQDLVERAAVRVLSGANPRAG 480
DB 467 RMDIDVTRTNKIRDRFWDPGPAADPLTDLYRWGGGFYLLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOQMPVPCVDDVFLRVLSRSLPLFLTLAWIYSVTITVKAVVREKETRLRDTMRMGL 540
DB 527 LYLOQMPVPCVDDVFLRVLSRSLPLFLTLAWIYSVTITVKAVVREKETRLRDTMRMGL 586
QY 541 SRAVLWLGWFLSCIGLPELSSAALLVLVLKLGDIILFYSHPGVVVFLFAFAVATVTSFLL 600
DB 587 SRAVLWLGWFLSCIGLPELSSAALLVLVLKLGDIILFYSHPGVVVFLFAFAVATVTSFLL 646
QY 601 SAPFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLDPAGGRVAASLLSPVAFGFCESIAL 660
DB 647 SAPFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLDPAGGRVAASLLSPVAFGFCESIAL 706
QY 661 LEEQEGCAQWNVCTRTADVFSLAQVSGLLLLDAALYGLATWLYEAVCPGOYGIPEPNW 720
DB 707 LEEQEGCAQWNVCTRTADVFSLAQVSGLLLLDAALYGLATWLYEAVCPGOYGIPEPNW 766
QY 721 PPFRRSYWCGRPPPKSPAPCPTPLDLPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRG 780
DB 767 PPFRRSYWCGRPPPKSPAPCPTPLDLPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRG 826
QY 781 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGGSAFTILGHNVRSMAAIRPHLVC 840
DB 827 LSLDFYQGHITAFILGHNGAGKTTTSLISGLFPPSGGSAFTILGHNVRSMAAIRPHLVC 886

```
QY 841 PQYNVLFDMITVDHWFYGRKLGSAVVGPEQDRLLQDVGLVSKSVOTRHLSGQMOR 900
Db 887 PQYNVLFDMITVDHWFYGRKLGSAVVGPEQDRLLQDVGLVSKSVOTRHLSGQMOR 946
QY 901 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDEAELLGD 960
Db 947 KLSVAIAFVGSGSVVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDEAELLGD 1006
QY 961 RVAVAGRLCCCGSPFLRRHLGSGYYLILVKARLPLTTNEKADTMEGSDVTRQSKN 1020
Db 1007 RVAVAGRLCCCGSPFLRRHLGSGYYLILVKARLPLTTNEKADTMEGSDVTRQSKN 1066
QY 1021 GSQSRVGTQQLALVQHWPGARLVEELPHELVLVLPYTCAGHDGSPATLFRLEDTELA 1080
Db 1067 GSQSRVGTQQLALVQHWPGARLVEELPHELVLVLPYTCAGHDGSPATLFRLEDTELA 1126
QY 1081 LRLTGYGISTSEIEIFLKVVEECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPPOETA 1140
Db 1127 LRLTGYGISTSEIEIFLKVVEECAADTDMEDGSCGQHLCTGTAGLDVTLRLKMPPOETA 1186
QY 1141 LENGEAGSAPETDQSGPDAVGVQWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1200
Db 1187 LENGEAGSAPETDQSGPDAVGVQWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1246
QY 1201 ALFVGLALVSLIYVPPGHYPALRLSPTMYCAQVSFESEADPGDPGRARLLEALLQEAGL 1260
Db 1247 ALFVGLALVSLIYVPPGHYPALRLSPTMYCAQVSFESEADPGDPGRARLLEALLQEAGL 1306
QY 1261 EEPVQHSRRFSAPEVAEPAEVAKVIASGNWTPSPSPACQSQPGARRLLPDCPAAGGP 1320
Db 1307 EEPVQHSRRFSAPEVAEPAEVAKVIASGNWTPSPSPACQSQPGARRLLPDCPAAGGP 1366
QY 1321 PPPQAVTSGSEVQNLTRNLSDFLVKTYPLRVQGLKTKKVNNEVRYGFSLGGRDPGL 1380
Db 1367 PPPQAVTSGSEVQNLTRNLSDFLVKTYPLRVQGLKTKKVNNEVRYGFSLGGRDPGL 1426
QY 1381 PSGELGRSEVELWALLSPLPGGALDRVLKNTAWAHSILDAQSLKTIWNNKGWHSWAF 1440
Db 1427 PSGELGRSEVELWALLSPLPGGALDRVLKNTAWAHSILDAQSLKTIWNNKGWHSWAF 1486
QY 1441 VNRSASNAIHLRAHLPFGPARIAHSITTLNHPNLNTKEQLESAALMASSVDVLVSIQVVFAM 1500
Db 1487 VNRSASNAIHLRAHLPFGPARIAHSITTLNHPNLNTKEQLESAALMASSVDVLVSIQVVFAM 1546
QY 1501 SFVPAFTLVLIERVTRAKHLQMGGLSPTLYWGNFLDMCMNLYVPACIVVLIFLAQ 1560
Db 1547 SFVPAFTLVLIERVTRAKHLQMGGLSPTLYWGNFLDMCMNLYVPACIVVLIFLAQ 1606
QY 1561 QRAYVAPANLPALLLLLLYCWSTPLMYPASFEFSPSTAYVYVLTCTINLFIGINGSMAT 1620
Db 1607 QRAYVAPANLPALLLLLLYCWSTPLMYPASFEFSPSTAYVYVLTCTINLFIGINGSMAT 1666
QY 1621 FVLELFSQDKLOEVSRIKQVFLIFPHFCLGRGLDMVRNQAMADAFERLGDROFQSPFLR 1680
Db 1667 FVLELFSQDKLOEVSRIKQVFLIFPHFCLGRGLDMVRNQAMADAFERLGDROFQSPFLR 1726
QY 1681 WEVVGNKLLAMVIOGPIFLFTLLQHSQOLLPOPRVRSPLLIGEEDEVARERVVQG 1740
Db 1727 WEVVGNKLLAMVIOGPIFLFTLLQHSQOLLPOPRVRSPLLIGEEDEVARERVVQG 1786
QY 1741 ATQGDVLVRLNLTKVYRGORMPADVRLCLGIPPGCEGFLGVNGAGKTSIFRWVTGDTLA 1800
Db 1787 ATQGDVLVRLNLTKVYRGORMPADVRLCLGIPPGCEGFLGVNGAGKTSIFRWVTGDTLA 1844
QY 1801 SRGEAVLAGHSVAREPSAHLMSG-YCPQSDAIEELL 1836
Db 1845 SFTHELLCIH-----HLLIGTYCM---PIFVLL 1869
```

RESULT 10

US-09-995-542-3

; Sequence 3, Application US/09995542

; Patent No. US20020127647A1

```
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Iarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2121
; TYPE: PRT
; ORGANISM: Mus musculus
; ORGANISM: Mus musculus
US-09-995-542-3
```

```
Query Match 77.0%; Score 8389.5; DB 9; Length 2121;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;

QY 1 PPLEHECHFPNKPFLPSAGTVPWLQGLICNNVNTCFPQLTPGPEGRGLSNFNDSLVRLL 60
Db 1 PPLEHECHFPNKPFLPSAGTVPWLQGLVCNVNNSCFQHPPTPEKPGVLSNFKDSLVRLL 60
QY 61 ADARTVLGGASAHRTTAGLGLKIATRAAKSTAQP-----OPTKQSPLEPPMLDVAELLTS 116
Db 61 ADRTVLGGHSIQDMLEDALGKLIPVLRVAGVGARPOESDQPTSQG-----SVTKLLEK 113
QY 117 LRTESIGLALGALGOAQLHSLLEAEADLAQELALASLVELRALLORPGTSGPLELLE 176
Db 114 ILORASLDVVLGOAQOSMRKFSDAIRDLAQELTLPSLMELRALLRRPGSASLELVSE 173
QY 177 ALCSVRGSPSTVGPSLNWYEASDLMELVGOEPESALPDSLSLSPACSELIGALDHPVRL 236
Db 174 ALCSVGRSPSGGLSNWYEANQNEFMGEVAPALPDNSLPACSEFVGTLDHPVRL 233
QY 237 LWRRLKPLILGLKLFAPDPTFTTKLMAQVNRTEETLLDREVWEMGLPRTFTMDS 296
Db 234 LWRRLKPLILGLKLFAPDPTFTTKLMAQVNRTEETLLDREVWEMGLPRTFTMDS 293
QY 297 SNVAMQLRLQOMDEGRORPRGGRDHMEALRGFLDPGSGGYSWODAHADVHLVGTGR 356
Db 294 TNVAMQLRLQOMDEGRORPRGGRDHMEALRGFLDPGSGGYSWODAHADVHLVGTGR 353
QY 357 VTCLSLDKLEAAPSAALVSRALQLLAEHRFWAGVVFLLGPEDSDPTTEHTPTDLGPGHV 416
Db 354 MMECVSLDKLEAVPSEALVSRALQLLAEHRFWAGVVFLLGPEDSDPTTEHTPTDLGPGHV 413
QY 417 RIKIRMDIDVVTNKRIRDFWDPGPAADPLTDLRVYWGFGFVYLODLVERAANVRVLSGAN 476
Db 414 RFKIRMDIDVVTNKRIRDFWDPGPAADPLTDLRVYWGFGFVYLODLVERAANVRVLSGAN 473
QY 477 PRAGLYLQMPYPCYVDDVFLVLSLPLFLTLAMIYSVTLTKAVVRKEKTRLRDTMR 536
Db 474 SRTGLYLQMPYPCYVDDVFLVLSLPLFLTLAMIYSVTLTKAVVRKEKTRLRDTMR 533
QY 537 AMGLSRAVLWGLWFLSCIGPFLLSAALLVLVLKLGIDILPYSHGFWVFLFAAFATVTTQ 596
Db 534 AMGLSRAVLWGLWFLSCIGPFLLSAALLVLVLKLGIDILPYSHGFWVFLFAAFATVTTQ 593
QY 597 SFLLSAPPSPANLAAACGGLAYPSLYLPIYVLCVAMDRPLPAGGRVAASLISPAVFGGCE 656
Db 594 SFLLSAPPSPANLAAACGGLAYPSLYLPIYVLCVAMDRPLPAGGRVAASLISPAVFGGCE 653
QY 657 SLALLEQEGGAQWHNVGTRPTADVFSLAQVSGLLLDALYGLATWYLEAVCPGOYGIP 716
Db 654 SLALLEQEGGAQWHNVGTRPTADVFSLAQVSGLLLDALYGLATWYLEAVCPGOYGIP 713
QY 717 EPWNFFRRSYWCGPRPPKSPAPCTPLDPAKVLVEEAPPGLSPGVSVRSLEKFPFGSPQP 776
Db 714 EPWNFFRRSYWCGPRPPKSPAPCTPLDPAKVLVEEAPPGLSPGVSVRSLEKFPFGSPQP 773
```

Qy	777	ALRGLSLDFVQGHITAFILGHNGAGKTTTILSLGLFPPSGSASFILGHVVRSSMAAIRPH	836
Db	774	ALQGLNLFDFEGHITAFILGHNGAGKTTTILSLGLFPPSGSASILGHVQTNMAAIRPH	833
Qy	837	LGVCQYNVLFDMLTVDBEHWFYGRLLKGLSAAVVGPDRLDODVGLVSKOSVQTRHLSG	896
Db	834	LGICQYNVLFDMLTVDBEHWFYGRLLKGVSAAMPREIRLIRDVGLTKRDTQTRHLSG	893
Qy	897	GMQRKLSVAIAFVGSGSQVILDEPTAGVDPASRRGIWELLIKYREGRTLILSTHHLDEAE	956
Db	894	GMQRKLSVAIAFVGSGSRVIMDEPTAGVDPASRRGIWELLIKYREGRTLILSTHHLDEAE	953
Qy	957	LLGDRVAVAGGRCCCGSPFLFRHLGSGYLLTLVKARLPLTTNE--KADTDMEGSDVDR	1015
Db	954	LLGDRVAVAGGSLCCCGSPFLFRHLGCGYLLTVKRSQSLVTHDAKGDSE----DPR	1008
Qy	1016	QEKKNQSGS-----RVGTQQLALVQHWPGARLVEELP	1050
Db	1009	REKSDNGRSTDTAFTRGTSDKSNQAPACGAVIPTSTARIILELVQHWPGARLVEDLP	1068
Qy	1051	HELVLVLPYTCAMDGSPATIFRELDTRLABELRTGYISDTSLEBIEFLKVVEECAADTDM	1110
Db	1069	HELLLVLPYAGALDGSFAMVFOELDQLELLGLTGYISDTNLEBIEFLKVVED--AHREG	1126
Qy	1111	EDSGCGHLCGTGIAGLDVTLRLKMPPOSTALENGEPAGSA--PETDQSGSPDVGVRQGVH	1168
Db	1127	GDGRPQJHLRT-----CTPQPTGPPEASVLENGELAKVLDPQAPQGLAPNA-AQVQGV	1179
Qy	1169	ALTTRQQLQALLKRFLLARSRGLFAQIVLPALFVGLALVFSLLVPPFGHYPALRLSP	1228
Db	1180	TLTCCQLRALLHKFLLARSRGLFAQVVLPAFVGLALFFSLVLPFPQYPPQLQSLPA	1239
Qy	1229	MYGAQVSFFSEDAPGDPGRARLLLEALLQEAAGLEBPPVQHSRHR-----FSAPEV	1277
Db	1240	MYGPQVSFFSEDAPGDPNRMKLLLEALLGEAGLQEPSMQDARGSECTHSLACYFTVPEV	1299
Qy	1278	PAEVAKVTLASGNWTPSPSPACCCSQPCARLLPDCPAAAGBPBPPOAVTGSSEVQNLT	1337
Db	1300	PPDVASITLASGNWTPSPSPACCCSQPCARLLPDCPAGAGBPBPPOAVAGLGEVQNLT	1359
Qy	1338	GRNLSDFLVKTYPRLVQRGLKTKWNVREYGGFSLGGRDPGLPSGQELGRSVEELWALL	1397
Db	1360	GRNVSDFLVKTYPSLVPRGLTKKWQDVEYGGFSLGGRDPDLPTGHEVVRTLAETALL	1419
Qy	1398	SPIPGGALDRVLKNLTAWAHSLDAQOSLKTWFNKNKGHSWVAFVNRASNAILRAHLP	1457
Db	1420	SPOGNAIDRLNLTQWALGLDARNLSKIWFNKNKGHWAVFVNRANGLLHALLPSGP	1479
Qy	1458	ARHAHSITTLNHPNLNLTKEOLSEAAWASSVDVLVSTCVVFAMSFVPASFTVLIEBVT	1517
Db	1480	VRHAHSITTLNHPNLNLTKEOLSEATLIASSVDVLVSTCVVFAMSFVPASFTVLIEBRT	1539
Qy	1518	RAKHQLMQGLSPITYLWGNFLMDMNCNLYVPACITVLILFLAFQORAYVAPANLPALLLLL	1577
Db	1540	RAKHQLVSLGPQTYLWGNFLMDMNCNLYVAVCIVVIFLAFQORAYVAPENLPALLLLL	1599
Qy	1578	LLYGWSITPLMPASFPFSPVSTAYVVLTCINLFIGINGSMATFVLELFSQDKLQEVRSI	1637
Db	1600	LLYGWSITPLMPASFPFSPVSTAYVVLTCINLFIGINSSMATFVLELLSDQNLQEVRSI	1659
Qy	1638	LKQVELIFPHFCILGRGLIDMVRNQAMADAFERLGDQFQSGPLREWYVGNKLAMWLOGPL	1697
Db	1660	LKQVFLIFPHFCILGRGLIDMVRNQAMADAFERLGDQFQSGPLRWDIIGKNLLAMWAGPL	1719
Qy	1698	FLLEFTLLQHRSQLLPQVRVSLPLGLEEDEDVARERERVQOGATQGDVLVLRLNLTKVYR	1757
Db	1720	FLLTITLLQHRNLLPQSKPRLPLPLEEDEDVAQREERVTKGATQGDVLVLRLDITKVYR	1779
Qy	1758	QORNPADRLICLIPGCECFGLLVNGAGKTSFRMWVTGDTLASRGEAVLAGHSVAREPS	1817
Db	1780	QORNPADRLICLIPGCECFGLLVNGAGKTSFRMWVTGDTLASRGEAVLAGHNVAQERS	1839

RESULT 11

US-09-995-542-2

; Sequence 2, Application US/09995542

; Patent No. US20020127647A1

GENERAL INFORMATION:

APPLICANT: Shutter, John

APPLICANT: Ulias, Laarni

TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and

; TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-658-A

; CURRENT APPLICATION NUMBER: US/09

; CURRENT FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 2000-

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Pat

; SEQ ID NO 2

; LENGTH: 2167

```

; TYPE: PRT

```

; ORGANISM: Mus musculus

Query Match 77.0%: Score 8389.5: DB 9: Length 2167:

Best Local Similarity 76.34: pred. No. 0: 3303.5

Matches 1636;	Conservative 161;	Mismatches 281;	Indels 65;	Gaps 11;
---------------	-------------------	-----------------	------------	----------

1 PPLEHHECHFPNKPLPSAGTVPWLOGLICNVNNTCFPOLTPGEEPGRLSNFNDSLVSRL 60

QY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
QY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00

37 FF DEETHECIT ENKEPDEFCAGIVFMDQDVCNVNNSCFQHFIFGEKFGVDSNFNSDLSKED 100

61 ADARTVLGSAHRTLGLKLIATLRAARSTAOP-----OPTKOSPLEPPMLDVAELLTS 116

Db 107 ADTRTVLGGHSIODMLDALGKLPVLRVGGGARPOESDOPTSO-----SVTKLEK 159

[illegible]

QY ERKES TO PATO QAF UNO DUE PAE DTA QED DUA KS DVE ENKAT QAF ROIS GF DE DUE I 17

db 160 ILORASLDPVLGOAODSMRKFSDAIRDLAQELLTLP SLMELRALRRRPRGSAGSLELVE 219

177 AI.CSVRGPSSTVGPST.NWVEASDI.MEIVGOEPESAI.PDSSI.SPACSEI.TGAI.DSHPI.SRI. 236

QY I / ATCSVRPSSIVGFSUNWIEASDLMELVQFEESATFPUSSLSFACSEFI GADSHFUSKT Z30

DD Z20 ATCSINGPSSFGGLSUNWIEANQLNEFMGPEVAFALPUNSLSPACSEFVGILDDHFFVSRU Z173

Qv 237 LWRRLKPLTIIGKLLFAPDTPETRKLMAQVNRTFFELTLRLDVREVVWMLGPRIETFMNDS 296

db 280 I.WRRI.KPI.II.GKTI.EAPD.TNETRKI.MAQVNOTFEEL.A.I.RDI.HEI.WGVI.GPOI.FNEMNDS 339

ZOO LMKKKLNF EIDGGK I D F A F D I N F I K N T W A Q V N Q I F B E S T W J R D P H E S M G V D S F Q I F N F M D S J C

Db 400 MMECVSLDKLEAVPSEALVSRLELGERRLWAGIVFLSPHEPLDPSELSPALSPCHL 459
Qy 417 RIKIRMDIDVTRNKNIRDRFWDGPAADPTDLRYVWGVFYLODLVERAAVRLSGAN 476
Db 460 RFKIRMDIDVTRNKNIRDRFWDGPAADPTDLRYVWGVFYLODLVERAAVRLSGAN 519
Qy 477 PRAGLYLOOMPYPVCVDDVFLVLSRSIPLFLTLAWIYSVTLTKVAVREKETRLDTMR 536
Db 520 SRTGLYLOOMPYPVCVDDVFLVLSRSIPLFLTLAWIYSVTLTKVAVREKETRLDTMR 579
Qy 537 AMGLSRAVLWGLFSLCIGPPLLSAALLVVLKGLDILPYSHPGWVFLFAAFVAVTQ 596
Db 580 AMGLSRAVLWGLFSLCIGPPLLSAALLVVLKGLDILPYSHPGWVFLFAAFVAVTQ 639
Qy 597 SFLLSAFPSRANLAAACGLAYFSLYLYPYLVCVNRDLRPAAGRVAASLLSPVAFGCE 656
Db 640 SFLLSAFPSRANLAAACGLAYFSLYLYPYLVCVNRDLRPAAGRVAASLLSPVAFGCE 699
Qy 657 SLALLEQGGCAOWHNVCTRTADVFSLAQSGLLLDAAALYGLATWYLEAVCGQYGP 716
Db 700 SLALLEQGGCAOWHNVCTRTADVFSLAQSGLLLDAAALYGLATWYLEAVCGQYGP 759
Qy 717 EPMNFPFRSRYWCGPFPKSPAPCTPLDPKVLVEEAPPGLSPGVSRSLKRPSPQP 776
Db 760 EPMNFPFRSRYWCGPFPKSPAPCTPLDPKVLVEEAPPGLSPGVSRSLKRPSPQP 819
Qy 777 ALRGLSLDFYOGHITAFIHLNGAGKTTLSLGLFPPSPGSAFILGHVSRSSMAAIRPH 836
Db 820 ALRGLSLDFYOGHITAFIHLNGAGKTTLSLGLFPPSPGSAFILGHVSRSSMAAIRPH 879
Qy 837 LGVCPQNVLFDMTLVDEHWFYGRKAGLSAAVGPEDQLLDQVGLVSKOSVOTRHLG 896
Db 880 LGVCPQNVLFDMTLVDEHWFYGRKAGLSAAVGPEDQLLDQVGLVSKOSVOTRHLG 939
Qy 897 GMQRKLSVAIAFVGSQVWILDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAE 956
Db 940 GMQRKLSVAIAFVGSQVWILDEPTAGVDPASRRGIWELLKYREGRTILSLTHLDEAE 999
Qy 957 LLGDRVAVAGRLCCGSPFLRRHLGSGYLLTLVKARLPLTNE-KAOTDMGSDVTR 1015
Db 1000 LLGDRVAVAGRLCCGSPFLRRHLGSGYLLTLVKARLPLTNE-KAOTDMGSDVTR 1054
Qy 1016 QEKKNQSGS-----RVGTPQLLALVQHWPGARLVEELP 1050
Db 1055 REKSDNGRTSDTAFTRTGSKNQAPAGVAPITESTARIELVQOHVPGAQVLEDLP 1114
Qy 1051 HELVLVLYTGADGSPATLFRDLRLAELRTGYGISTSLSEIEFLKVVECAADTM 1110
Db 1115 HELVLVLYTGADGSPATLFRDLRLAELRTGYGISTSLSEIEFLKVVECAADTM 1172
Qy 1111 EDGSCGHLCTGIAGLDVTLRLKMPPOETALENGEPAGSA--PETDQSGPDVAVRQVGH 1168
Db 1173 GDSRPLQLHRT-----CTPQPTGPBASVLENGELAKLVLDPOAQGLAPNA-AQVQGW 1225
Qy 1169 ALTRQQLQALLKRLFLARSRLGFAQIVLPALFVGLALVESLIVPPFGHYPALRLSPT 1228
Db 1226 TLTCQQRALLKRLFLARSRLGFAQIVLPALFVGLALVESLIVPPFGHYPALRLSPT 1285
Qy 1229 MYGAQVFFSDEADPGDPRARLLLEALQEALEPPVQHSR-----FSAPEV 1277
Db 1286 MYGAQVFFSDEADPGDPRARLLLEALQEALEPPVQHSR-----FSAPEV 1345
Qy 1278 PAEVAKVLASGNWTPESPACQSQCARLLPDCPAAAGGPPPOAVTGSVEVONLT 1337
Db 1346 PAEVAKVLASGNWTPESPACQSQCARLLPDCPAAAGGPPPOAVTGSVEVONLT 1405
Qy 1338 GRNLSDFLVKTYPRVQGLTKKWNNEVRVYGFSLGCRDPLPSGOELGRSVEELWALL 1397
Db 1406 GRNLSDFLVKTYPRVQGLTKKWNNEVRVYGFSLGCRDPLPSGOELGRSVEELWALL 1465
Qy 1398 SPLPGGALDRVLKNTAWAHSLSAQDSLKIFWNNKGMHSMVAFVNRASNAIRLAHLPGP 1457
Db 1466 SPLPGGALDRVLKNTAWAHSLSAQDSLKIFWNNKGMHSMVAFVNRASNAIRLAHLPGP 1525

Qy 1458 ARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVIVPAMSFVPASFTLVLIBERV 1517
Db 1526 VRHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVIVPAMSFVPASFTLVLIBERV 1585
Qy 1518 RAKHLQMLMGGISPTLWGLNFMNLCNLYLPACIVVLI FLAFOORAVVAPANLPALLLLL 1577
Db 1586 RAKHLQMLMGGISPTLWGLNFMNLCNLYLPACIVVLI FLAFOORAVVAPANLPALLLLL 1645
Qy 1578 LLYXGWSITPLMYSPASFFSVSTAYVVLTCINIFIGINGSMATFVLELSDQKLEYSRI 1637
Db 1646 LLYXGWSITPLMYSPASFFSVSTAYVVLTCINIFIGINGSMATFVLELSDQKLEYSRI 1705
Qy 1638 LKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDQFQPSLRWVVGKLLAMVIOGPL 1697
Db 1706 LKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDQFQPSLRWVVGKLLAMVIOGPL 1765
Qy 1698 FLITLTLLOHRSOLLPOQVRSLPLLGEDEDAVABRERVVOGATOGDVLVLRNLTQVYR 1757
Db 1766 FLITLTLLOHRSOLLPOQVRSLPLLGEDEDAVABRERVVOGATOGDVLVLRNLTQVYR 1825
Qy 1758 GQMPAVDRCLCLGIPPECECFGLLGVNGAGKTSIFRMVVTGDTLASRGEAVLAGHSVAREPS 1817
Db 1826 GQMPAVDRCLCLGIPPECECFGLLGVNGAGKTSIFRMVVTGDTLASRGEAVLAGHSVAREPS 1885
Qy 1818 AAHLGWSGCPQSDAIFELITGRHLELLARLGRVPEAQAQTAGSGLARLGLSWYADRP 1877
Db 1886 AAHLGWSGCPQSDAIFELITGRHLELLARLGRVPEAQAQTAGSGLARLGLSWYADRP 1945
Qy 1878 GTYSGGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLMNSLLAVVREGSVMLTSH 1937
Db 1946 GTYSGGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLMNSLLAVVREGSVMLTSH 2005
Qy 1938 SMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAPFAAEFP 1997
Db 2006 SMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAPFAAEFP 2065
Qy 1998 GSELREAHGRLRQLPQGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYPSKD 2057
Db 2066 GSELREAHGRLRQLPQGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYPSKD 2125
Qy 2058 QGKDEDETEOKKAGVGDPAQGLQHPKRVSOFLDDPSTAEVTL 2100
Db 2126 QGKDEDETEOKKAGVGDPAQGLQHPKRVSOFLDDPSTAEVTL 2167

RESULT 12

US-09-995-542-8
; Sequence 8, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1550
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-8

Query Match 69.7%; Score 7596; DB 9; Length 1550;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 121 ESIGLALGQAQEPHLSLEAAEDLAQELLALRSVLRLALQRPRTSGPFLLEALCS 180

Db 29 ESIGLALGQAQEPHLSLEAAGDQAQELLARSLVELRALLQRPCTSGPLELLSEALCS 88
Qy 181 VEGBSSTWCPSPNWEASDLMEIVGQEPESALPDASLSLSPACSELICALDHPHLSLLWRR 240
Db 89 VEGPSTWCPSPNWEASDLMEIVGQEPESALPDASLSLSPACSELICALDHPHLSLLWRR 148
Qy 241 LKPLILGKLLFAPDPTFTTKMAQVNRTPPEELTLRDVREVMELGPRIFTFMNDSSNVA 300
Db 149 LKPLILGKLLFAPDPTFTTKMAQVNRTPPEELTLRDVREVMELGPRIFTFMNDSSNVA 208
Qy 301 MLQRLLOMDQDEGRQPRPGGRDHMBALRSFLDPGSGGYSWQDAHDVGHVGLVGTGRVTEC 360
Db 209 MLQRLLOMDQDEGRQPRPGGRDHMBALRSFLDPGSGGYSWQDAHDVGHVGLVGTGRVTEC 268
Qy 361 LSLDKLEAPSAALVSRALQLLAHRFWAGVVFILGPEDSSDPTHEPTDPLGPGHVRKI 420
Db 269 LSLDKLEAPSAALVSRALQLLAHRFWAGVVFILGPEDSSDPTHEPTDPLGPGHVRKI 328
Qy 421 RMDIDVTRTKIRORFMDPGPAADPLTDLRYVMGFGVYLQDLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRORFMDPGPAADPLTDLRYVMGFGVYLQDLVERAAVRVLSGANPRAG 388
Qy 481 LYLQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 540
Db 389 LYLQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL 448
Qy 541 SRAVLWLGWFLSLCLGPFLLSALLVLVLKGLDILPYSHPGVVFLLAFAFAVATVQSFL 600
Db 449 SRAVLWLGWFLSLCLGPFLLSALLVLVLKGLDILPYSHPGVVFLLAFAFAVATVQSFL 508
Qy 601 SAFFSRANLAAACGLAYFSLYLPVVLVAVNRDLPAAGRVAASLLSPVAFGFCESLAL 660
Db 509 SAFFSRANLAAACGLAYFSLYLPVVLVAVNRDLPAAGRVAASLLSPVAFGFCESLAL 568
Qy 661 LEEQEGQAQHNVTGRTADVFLSAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPN 720
Db 569 LEEQEGQAQHNVTGRTADVFLSAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPN 628
Qy 721 PPFRRSYWCGRPKPSAPCPPTLPDKVLVEAPPGSLPGSVRSLEKRRPGSPQALRG 780
Db 629 PPFRRSYWCGRPKPSAPCPPTLPDKVLVEAPPGSLPGSVRSLEKRRPGSPQALRG 688
Qy 781 LSLDFYQCHIYAFIHLGNAGKTTLSILSGILPPSPGSAFILGHDRVSSMAAIRPHLVC 840
Db 689 LSLDFYQCHIYAFIHLGNAGKTTLSILSGILPPSPGSAFILGHDRVSSMAAIRPHLVC 748
Qy 841 POYNVLFDMLTVDHWFYGRILKGLSAAVVGPEQDRLLQDVLGSKSVQTRHLSGGMR 900
Db 749 POYNVLFDMLTVDHWFYGRILKGLSAAVVGPEQDRLLQDVLGSKSVQTRHLSGGMR 808
Qy 901 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKXYREGRTLIILSTHLDRAELLGD 960
Db 809 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKXYREGRTLIILSTHLDRAELLGD 868
Qy 961 RVAVVAGRLCCGSPFLRHLGSGYTLVAKRLPLTTNEKADTDMEGSVDRQEKKN 1020
Db 869 RVAVVAGRLCCGSPFLRHLGSGYTLVAKRLPLTTNEKADTDMEGSVDRQEKKN 928
Qy 1021 GSQSRVGTQPOLALVQHWVFCARLVEELPHELVLVLPYTGANDGSPATLFRDLTRLAE 1080
Db 929 GSQSRVGTQPOLALVQHWVFCARLVEELPHELVLVLPYTGANDGSPATLFRDLTRLAE 988
Qy 1081 LRLTYGIGSDTSLBIEFLKVVVECAADTDMEDGCGHLCGTGAGLDVTLRLKMPPOSTA 1140
Db 989 LRLTYGIGSDTSLBIEFLKVVVECAADTDMEDGCGHLCGTGAGLDVTLRLKMPPOSTA 1048
Qy 1141 LENGEPAQAPETDGGSPDAGVRVQGWALTRQQLQALLKRFLLARSREGLFAQIVLP 1200
Db 1049 LENGEPAQAPETDGGSPDAGVRVQGWALTRQQLQALLKRFLLARSREGLFAQIVLP 1108
Qy 1201 ALFVGLALVFSILVPPFGHPALRLSPMTYGAQVFSFFSEDAFGDPGRARLLEALLQEA 1260

Db 1109 ALFVGLALVFSILVPPFGHPALRLSPMTYGAQVFSFFSEDAFGDPGRARLLEALLQEA 1168
Qy 1261 BEPPVQHSRPSADPEVAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1320
Db 1169 BEPPVQHSRPSADPEVAEVAKVILASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1228
Qy 1321 PPPQAVTSGGEVQVQNLGTNRISDFLVKTYPRLVROGLTKTKWVNEVRYGGFSLGGDRDGL 1380
Db 1229 PPPQAVTSGGEVQVQNLGTNRISDFLVKTYPRLVROGLTKTKWVNEVRYGGFSLGGDRDGL 1288
Qy 1381 PSGELGRSVEELWALLSPLPGALDRVLKNLTAWAHSIDLAQDSLKIWNKNGHSMVAF 1440
Db 1289 PSGELGRSVEELWALLSPLPGALDRVLKNLTAWAHSIDLAQDSLKIWNKNGHSMVAF 1348
Qy 1441 VNRSNATILRAHLPPGPARHAHSITTLNHLPLNLTKEQLSEALMASSVDVLVSI 1500
Db 1349 VNRSNATILRAHLPPGPARHAHSITTLNHLPLNLTKEQLSEALMASSVDVLVSI 1408
Qy 1501 SFVPASFTVLIEBRVTRAKHLQLMGGLSPTLYMLGNFLMDMNCNLYLVPA 1560
Db 1409 SFVPASFTVLIEBRVTRAKHLQLMGGLSPTLYMLGNFLMDMNCNLYLVPA 1468
Qy 1561 QRAYVAPANLPALLLLLYGWSITPL 1587
Db 1469 QRAYVAPANLPALLLLLYGWSITPL 1495

RESULT 13

US-10-313-641-9
; Sequence 9, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P02351US2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
; US-10-313-641-9

Query Match 51.5%; Score 5610.5; DB 14; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
Qy 1 PPLEHHECHFPNKLPSAGTVPMLOGLICNVNNTCFPOLTFGBEPGRLSNFDLSVRL 60
Db 47 PPVEQHECHFPNKAAPSAGTLPWYQGIICNANNPCFRYPTEGAPGVVGNFKSIVARLF 106
Qy 61 ADARTVLGGASAHRTLAGLKLATLRAARSTAQ-----DVAEL-----PQ 96
Db 107 SDARLLLYSQDTSKMDKRVLTLOQIKKSSNLKLDQFLVDNETPSGFLYHNLSLPK 166
Qy 97 PTKQSPLEPPML-----DVAEL----- 113
Db 167 STVDKMLRADVLHKVFLQYQLHLTSLCNGSKSEMIQLGDQEVSELGCPREKLAAB 226
Qy 114 -----LTSLRT--BSLGLALGQAQEPHLSLEAAEDLAQELLALRSVLRL-- 159
Db 227 RVLRSNMDILKPLIRLTINSTSPPSKELAEATKTLHLSGLTLAGELFMRSDMRQEV 286

APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P023151US2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2261
; TYPE: PRF
; ORGANISM: Human
US-10-313-641-10

Query Match 51.5%; Score 5610.5; DB 14; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
QY 1 PPLEHHECHFNKPLPSAGTVPWLOGLICNVNNTCFPOLTPGEEPRGLSNFNDLSIVSELL 60
DB 47 PPEOHECHFNKAMPASAGTLPWQGIICNANNPCFRYPPTGEPAGVGVGNFNKSIARLF 106
QY 61 ADARTVLGGAHARTLAGLGLIATLRAARSTAO-----PQ 96
DB 107 SDARRLLYSQKDTSMKMDKRVLTLOQIKSSSNKLQDFLVNETPSGFLYHNLSLPK 166
QY 97 PTKOSPLBPPML-----DVAEL----- 113
DB 167 STVDKMLRADVLHKVFLQYQLHLTSLCNGSKSEMIQLGDQEVSESLCGLPREKLAAGE 226
QY 114 -----LTSLLRT--ESGLALGOAQOEPHLSLEAAEDLAQELLALRLSIVELR--A 159
DB 227 RVLRSNMDILKPIRLTNSTPSPSKELAEATKLLSHSLGLTLAGELFSMRSDMRQEVW 286
QY 160 LLQRPRTGSGPLEL---LSEALCSVRGSSVTGSLNWEASDLMELVG---OEPESAL 212
DB 287 FLTNVSSSSSTQIYQAVSRVICHGPEGGLKIKSLNWEYDNNYKALPGGNGTDEATFP 346
QY 213 PDSSLSACSELIGALDSHPISRLRLKPLILGLKLPADPTTPTRKLMQAQVNRTEEL 272
DB 347 YDNSTTPYCNLKNLESSPLSRIIWALKPLLVGKILYTPDTPATQVMAEVNKTFOEL 406
QY 273 TLLRDVREVMELGPRIETFNDSNVAMLRQLLOMQDBGR-RQRPQGRD----- 322
DB 407 AVFHDLEGMEELSPKWTFTFENSQEMDLVRLMLDSRDNDHFWSQQLDGLDWTQADIVAF 466
QY 323 ---HMEALRSLDPGSGG--YSWQDAHADVGHVLTGLGRVTECLSLDKLEAAPSEALVS 377
DB 467 LAKHPEDVQS-----SNGSVYTWREAFNETQARTISRFEVCNVLNKLPIATEVWLN 521
QY 378 RALQALLAHRFWAGVFLGPBDSDDTEHPDPLGPGHVRKIRMDIDVVTNKRIDRF 437
DB 522 KSMELLDERKFWAGVFTGITPGSTELPH-----HVKKIRMDIDNVERTNKKIDGY 573
QY 438 WDRGPAADPLTDLRYVWGGFFVLODVERAAVRVLSGANPRAGILYLOQMPYPCYVDDVFL 497
DB 574 WDPGRADPFEDMYRVWGGFAYLODVVEQAIIRVLTGTEKKTGVYMQQMPYPCYVDDIFL 633
QY 498 RVLGRSLPLFTLTAWIYSVTLTVKAVVREKETRLDRTRAMGLGRAVLWLGWFLSLCLGPF 557
DB 634 RVMSRSLPFTLTAWISYVAIIKGIYVEKARLKETWRINGLONSILWFSWFISSLIPL 693
QY 558 LLSAALLVLKLGIDILPSYHGVVFLPLAAFAVATVQSFLLSAFFSRANLAAACGLA 617
DB 694 LVSAGLLVILKLGILLPSYDPSVVFVFLSVFAVVTILQCFLISTLFSRANLAAACGII 753
QY 618 YFSLYLPVLCVAVNRDLRPAAGRVAASLLSPVAFGCGESLALLEEQEGEGQWNVGTRP 677
DB 754 YFTLYLPVLCVAVQYVGFLLKIFASILLSPVAFGCGEYFALFEPGQIGVQWDLNLFESP 813

QY 678 T-ADVFSLAQVSGLLLLDAAALYGLATWYLEAVPCQGYGIPBEPNPPRRSYWCGRPBPKS 736
DB 814 VEEDGFNLTVSSMMLFDTFLYGVWWTWYIEAVFPQGYGIPRPWYFPCTKSYWFGESDEK 873
QY 737 PAPCPTPLD-PKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRGLSLDFYQGHITAFILG 795
DB 874 SHPGSNOKRISEICWEEBPTHKLGVSIGNLVKVVYRDMKVAVDGLALNFYEGQITSFLG 933
QY 796 HNGAGKTTLSILSGLPPSPGSAFILGHVYRSSMAAIRPHLGYCPCQNVNLFDMLTVDSEH 855
DB 934 HNGAGKTTLSILSGLPPSPGSAFILGHVYRSSMAAIRPHLGYCPCQNVNLFDMLTVDSEH 993
QY 856 VWFYGRKLGLSAAVVGPEQDRLLODVLG-VSKQSVQTRHLSGGMQRKLSVAIAFVGGSGV 914
DB 994 IMFYARLKLSEKHVKAEMEOMALDVLGPPSKLSKTSQLSGGMQRKLSVALAFVGGSKV 1053
QY 915 VILDEPTAGVDPASRRGIWELLKLYREGRTIILSTHLDDEALLGDRVAVVAGRLCCCG 974
DB 1054 VILDEPTAGVDPASRRGIWELLKLYREGRTIILSTHLDDEALLGDRVAVVAGRLCCCG 1113
QY 975 SPLFLRRHLSGYYLTLVKARLPLTTN-----EKADTDMEGSVDTROEKNGS 1022
DB 1114 SSLFLKNQLGTYLTLVKDVSSESSLSRNSSTSVLYLKKEDSVSQSSDAGLSDBES 1173
QY 1023 QGSRVGTPLIALVQHWPVGARLVEELPHELVLVLPYTGADHSGFATLFRLEDTRLAELR 1082
DB 1174 DTLTIDVSAISNLKRVHSEARLVEDICHELTYVLPYEAKEGAFVELFHEIDRLSDLG 1233
QY 1083 LTGYGIDSLSLEIPLKVEECAADTMDSCGQHLCTGTIAGLDVTLRLKXMPQOETALE 1142
DB 1234 ISSYGISETTEELFLKVAESGVDAETSOGTLPARRNRRAFG-DKQSCLPFTTEDDAAD 1292
QY 1143 --NGEPAAGSAPETDQSGSDPAVG--RVQGNWALTQOQLQALLKFLARRRGRGLFAQIV 1198
DB 1293 PNDSDIDPESRETDLSSMDGKGSYQVKGWKLTOQFVALLWKKLLIARRSRKGFQFAIV 1352
QY 1199 LPALFVGLALVFSILVPPFGHYPALRLSPMTMYGAQVSPFSEDAQDPCGRARLLALQEA 1258
DB 1353 LPAVFVCIALVFSILVPPFGHYPALRLSPMTMYGAQVSPFSEDAQDPCGRARLLALQEA 1412
QY 1259 G-----LEBPPVQHSRHSFAPEVPAEVAKVLAASGNWTPESPSPACQSQPGARR 1308
DB 1413 GFGTRCMGNPIPTPCQAGEEEMTAPVQTIIMDLFQNGNWTWQNPSPACQSSDKIKK 1472
QY 1309 LLPDCPAAAGPPPPQAVTGSGEVQNLGTGNSLDFLVKTYPLRVROGLTKKWNVEVRY 1368
DB 1473 MLPVCPGAGGLPPQQRKQNTADILQDLTGNISDYLVKTYVQIIAKSKNKIWNVEVRY 1532
QY 1369 GGFSLG-GRDPGLPSGQELGRSVLEELWALLSPLPGGALDRVLKNLTAWAHSLDAQDSLKI 1427
DB 1533 GGFSLGVSNTQALUPPSQEVNDIAIKOMKKHLKLAQDSSADRFNLNSLGRFMTGLDTKQNVKV 1592
QY 1428 WFNKKGHSMVAFVNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTQKQSEALMASS 1487
DB 1593 WFNKKGHSMVAFVNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTQKQSEALMASS 1652
QY 1488 VDLVLSICVVFAMSPVPSFTLVLEIERTVTRAKHLMGGLSPTLYMLGNFLDMCNVLA 1547
DB 1553 VDLVLSICVVFAMSPVPSFTLVLEIERTVTRAKHLMGGLSPTLYMLGNFLDMCNVLA 1712
QY 1548 PACIVLIFLAFQORAVAPANLPAALLLLLLYGSITPLMYSPFFSFPSTAYVVLTC 1607
DB 1713 PATLVIIIFCFQOKSVSVSTNLPVALLLLLLYGSITPLMYSPFFSFPSTAYVVLTC 1772
QY 1608 INLFIGINGSMATFVLEFSDQKQVSRILKQVFLIPFHPFCIGRGLIDWVRNQMADAF 1667
DB 1773 VNLFIGINGSVATFVLEFSDQKQVSRILKQVFLIPFHPFCIGRGLIDWVRNQMADAF 1832
QY 1668 ERLGDRQFQSLRNEVVGKLLAMVIOGFLFLLFTLLQHSRQLLQPRVRSPLIGSED 1727
DB 1833 ERFGENFRVPSLDVGRNLFAMAVEGVVFFLTIVLIQVRRFFIRPRFVNAKLSPLNDED 1892
QY 1728 EDVARERVRVQCATQGDVLVRLNLTKVYRGQMPAVDRCLCLGIPPGECFGLLGVNGAGK 1787

Db 1893 EDVREKORILDGGQNDILEIKELTKIYRKRKPAVDRCVGLPPGECFGLGVNGAGK 1952
Qy 1788 TSFPMVTGDTLASRGAVALAGHSAVAREPSAAHLSMGYCQSDAIFELLTGREHLELLAR 1847
Db 1953 SSFPMKLTGDTTVTRGDAFLNKSILSNIHEVHONMGYCQFQDAITELLTGREHVEFFAL 2012
Qy 1848 LRGVPEAQVTAQTAGSLARGLSWYADRPACTYSGGNKRKLATLALVGDPAVVFLDEPT 1907
Db 2013 LRGVPEKGVKGWEAIVKGLGVKGYKAGNYSGGNKRKLSTAMALIGPPVVFLEPT 2072
Qy 1908 TGMDDPSARRFLWLSLAVREGSRVMLTSHSMEECEALCSRLATMNGRCULGSPHLLK 1967
Db 2073 TGMDDPKARRFLMNCALSVKREGSRVMLTSHSMEECEALCTRMAIMVNGRFRCLGSOHLK 2132
Qy 1968 GRFAAGHTLRLVPAARS--QPAFAAFAVAAFFPGSELREAGGRRLRFOLPPGRCALARVF 2025
Db 2133 NRFGDGYTVIRIAGSNPDLKPVODFFGLAFPGSVLKEKHNMLQYOL--PSSLSLARIF 2191
Qy 2026 GELAVHGAEGHVEDFVSQTMLEEVFLYFKDQKOE---DTEBQKEAGVGVDPAQGLQH 2082
Db 2192 SILSQSKRLHIEDYSVSQTLDOQVFNFAKQSDDDHLKDLSLHKNQTV--VDVAV---- 2246
Qy 2083 PKRVSQFLDDPSTAET 2098
Db 2247 ---LTSFLQDEKVKES 2259

RESULT 15
US-10-428-551-9
; Sequence 9, Application US/10428551
; Publication No. US20030229062A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P0231US3
; CURRENT APPLICATION NUMBER: US/10/428,551
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-428-551-9

Query Match 51.58; Score 5610.5; DB 14; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
Qy 1 PPLEHHECHFPNKPPLPSAGTVPMLOGLICNNVNTCFPOLTPGEPGRLSNFDNLSVRL 60
Db 47 PPEHHECHFPNKPMPAGTLPWVQGLICNNANPCFRYPPTGPAVGVGNFNKSIARLF 106
Qy 61 ADARTVLGGSASRHTLAGLKLATLRAASTAQ-----PQ 96
Db 107 SDARRLLYSOKOTSMKMRKVLRTLOQIKKSSNKLQDLVDNETFSGFLYHNLSLPK 166
Qy 97 PTKQSPLEPPML-----DVAEI----- 113
Db 167 STVDKMLRADVILHKVFLQYQLHLTSLCNGSKSEEMIQLDQOEVSLEGLPREKLAEE 226
Qy 114 -----LTSILRT--ESIGLALGOAQPFLSHLEAAEDLAQELIARLSVELR---A 159
Db 227 RVLRSNMILKPIRLTINSTSPFSPKELAEATKLLHSLGLTLAGELFSMRSWSDMRQEV 286

Qy 160 LLQRPKGTSGPLEL---LSEALCSVRGSPSSTVGPNSLWNYEASDLMELVG---QEPESAL 212
Db 287 FLTNVSSSSTOIYQAVSRIVCGHPGGGLKIKSLNWNEDNNYKALFGNGTEEDAETF 346
Qy 213 PDSLSPACSELIGALDSDHPLSLLWRLKPLILGKLIFAPDTPFTRKLAQVNRTEEL 272
Db 347 YDNSTTPYCNLDKMNLESSPLSRIIWKALKPLLVGKILYTPDTPATQVMAEVNKTFOEL 406
Qy 273 TLLRDVREVMELGPRIFTFMDNSNVMLQRLQWODEGR--RQRPGGRD----- 322
Db 407 AVFHDLEGMWEEELSPKLTWFMENSQEMDLVRLMLDSRDNDHFWEOQLDGLDWAQDIVAF 466
Qy 323 ---HMEALRSFLDPGGGG--YSWODAHADVGHVLTGRVTECLSDKLXAAASEAALVS 377
Db 467 LAKHPEDVQS-----SNGSVYTWEAFNETNQAITRISRFMECVNLNKLPIATEVWLN 521
Qy 378 RALOLLAERHFWAGVFLGPEDDSDTEHPTDLPDGHVRIKIRMDIDVVTNKNKIDRF 437
Db 522 KSMELDDEKFWAGIVFTGITPGSIELPH-----HVYKIRMDIDNVERTNKIKDGY 573
Qy 438 WDPGPAADPLTDLRYVMGFGVYLQDLVERAAVRLSGANPRAGLYLQOOMEYPCYVDVFL 497
Db 574 WDPGPRADPEDMRYVMGFGAYLQDVVEQAIIRVLGTGKTKGVYMOOMEYPCYVDIFL 633
Qy 498 RVLRSPLPLFLTLAWIYSVTLTKVAVVREKETRLDRTMRAMGLSRVAVLWGLFSLCLGPF 557
Db 634 RVMSRSMPLFMTLAWIYSVAVIKGIYVEKEARLKETMRIMGDLNDSNLFWSFISSLIPL 693
Qy 558 LLSAALLVLKGLDILPYSHPGVPLFLAAFAVATVQSFLLSAFESRANLAACGLA 617
Db 694 LVSAGLLVWLKLGNNLLPYSDPSVVFVLSVFAVVTILOCFLLSTLFSRANLAACGGII 753
Qy 618 YFSLYLPYVLCVARDRLPAGRVAASLLSPVAFGFCESLALLEEGEGEAQWNVTRP 677
Db 754 YFTLYLPYVLCVARDVGVFTLKFASLLSPVAFGFCGYFALFEQIGVQWDLNLPSP 813
Qy 678 T-ADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGQYQVPEPWNFFPRRYSWYCGPRPKS 736
Db 814 VEEDGFLNLTSSVMMLFDTFLYGVMTWYIEAVFPQYQVIGIPRPYFCTKSYNFESEDEK 873
Qy 737 PAPCTPLD--PKVLVEEAPGLSPGVSRLEKRFPGSOPALRGLSLDYPQGHITAFGL 795
Db 874 SHEGSKRISEICMEEEPHLKLGVSIOVLKYRDKMKVAVDGLALNFYEGQITSFLG 933
Qy 796 HNGAGKTTTILSILGFLPPSGSAFILGHVDVRSMAAIRPHLGVCPQYVNLFDMLTVDEH 855
Db 934 HNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPCQHNVLFDMLTVDEH 993
Qy 856 WIFYGRKGLSAAVGPEDORLLQDVGL--VSKOSVOTRHLSCGMQRKLSVAIAFVGGSQV 914
Db 994 IWFYARLKGLSEKHVKAEMEQMALDVLGFLPSKLSKTSQLSGGMQRKLSVALAFVGGSKV 1053
Qy 915 VILDEPTAGVDPASRRGIWELLILKYREGRTLILSTHLLDEALLGDRVAVVAGRLCCG 974
Db 1054 VILDEPTAGVDVPSRRGIWELLILKYRQRTIILSTHMDADVIGDKRIIISHLKCCVG 1113
Qy 975 SPLFLRRHLGSGYYLTLVKARLPLTTN-----EKADTDMEGSVDTROEKNGS 1022
Db 1114 SSLFLKNQLGTGYLTLVKDVESLSSCRNSSSTVYLKCKEDSVSSQSSDAGLGDHES 1173
Qy 1023 QGSRVGTPLLALVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLRELDTRLAELR 1082
Db 1174 DTLTI DVSAISNIRKHVSEARLVEDIGHELTYLPLPEAAKEGAFVELFHEIDRLSDLG 1233
Qy 1083 LTCYGDISTLSLEIFLKVVEECAADTDMEDSGQHLCTGCIAGLDVTLRLKMPQETALE 1142
Db 1234 ISSYGLSEITLLEIFLKVABESVDATSDGTLPARRNRFAFG--DKQSLRPTTEDAAD 1292
Qy 1143 --NGEPAQSAPETDOGSGPDVAG--RVQGWALTQOQLALLKRLFLARRSRRLGFAQIV 1198
Db 1293 PNDSDIDPESRETDLLSGMDGKGSYQVKGWKLTOQQFVALLWKRLIARRSRKGFQIV 1352

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 177.061 Seconds
(without alignments)
4254.650 Million cell updates/sec

Title: US-09-995-542-6
Perfect score: 10896
Sequence: 1 PPLEHHEFFNKPPLPSAGT.....QHPKRVSFQLDPPSTAETVL 2100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10896	100.0	2146	6	Abg72696 Human ATP
2	10883	99.9	2146	4	Aau04483 Human PD-
3	10883	99.9	2180	5	Aaol4210 Human tra
4	10881	99.9	2146	5	Abp52096 Homo sapi
5	10881	99.9	2146	6	Abu08466 Human ABC
6	10880	99.9	2144	5	Aau09174 Human tra
7	10880	99.9	2144	7	Add37429 Human tra
8	10392.5	95.4	2059	6	Abu54629 Human NOV
9	10264	94.2	2008	6	Abu08464 Amino aci
10	10171.5	93.4	1993	6	Abu08465 Amino aci
11	9213.5	84.6	1873	4	Aau04484 Human PD-
12	8389.5	77.0	2167	6	Abg72695 Mouse ATP
13	7596	69.7	1550	6	Abg72697 Human ATP
14	5611.5	51.5	2261	3	Aab38111 Human ABC
15	5611.5	51.5	2261	3	Aab38114 Human ABC
16	5610.5	51.5	2261	3	Aab38117 Human ABC
17	5610.5	51.5	2261	3	Aab38115 Human ABC
18	5610.5	51.5	2261	3	Aab38109 Human ABC
19	5610.5	51.5	2261	3	Aab38082 Human ABC
20	5610.5	51.5	2261	3	Aab38112 Human ABC
21	5610.5	51.5	2261	4	Aab71749 Human ABC
22	5610.5	51.5	2261	4	Aab31361 Amino aci
23	5610.5	51.5	2261	4	Aab31365 Amino aci
24	5610.5	51.5	2261	6	Abu11899 Human ATP
25	5610.5	51.5	2261	6	Abg72696 Human ABC

26	5610.5	51.5	2261	6	ABR62033	Humun ABC
27	5610.5	51.5	2261	7	ADP65173	Humun ATP
28	5609.5	51.5	2261	3	AAB38116	Human ABC
29	5608.5	51.5	2261	3	AAB38113	Human ABC
30	5606.5	51.5	2261	3	AAB38110	Human ABC
31	5606.5	51.5	2261	3	AAB38105	Human ABC
32	5606.5	51.5	2261	4	AAB31362	Amino aci
33	5606.5	51.5	2261	4	AAB31366	Amino aci
34	5602.5	51.4	2261	4	AAB31363	Amino aci
35	5602.5	51.4	2261	4	AAB31367	Amino aci
36	5600	51.4	2260	3	AAB38106	Human ABC
37	5598.5	51.4	2261	3	AAB38104	Human ABC
38	5594.5	51.3	2261	5	ABB83120	Polymorph
39	5594.5	51.3	2261	5	ABB83116	Polymorph
40	5594.5	51.3	2261	5	ABB83124	Polymorph
41	5594.5	51.3	2261	5	ABB83122	Polymorph
42	5593.5	51.3	2259	3	AAB38107	Human ABC
43	5593.5	51.3	2261	4	AAM50228	Human ATP
44	5593.5	51.3	2261	4	AAB13022	Human ATP
45	5593.5	51.3	2261	4	AAU02176	Human ABC

ALIGNMENTS

RESULT 1
ABG72696
ID ABG72696 standard; protein; 2146 AA.
XX AC ABG72696;
XX DT 10-MAR-2003 (first entry)
XX DE Human ATP-binding cassette transporter-like protein, ABCL.
XX KW Human; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukemia; neutropenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX OS Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..46
XX Protein /label= Signal_peptide
XX /label= Mature_ABCL
XX US2002127647-A1.
XX PD 12-SEP-2002.
XX PF 28-NOV-2001; 2001US-00995542.
XX PR 28-NOV-2000; 2000US-0253520P.
XX (SHUT/) SHUTTER J.
XX (ULIA/) ULIAS L.
XX Shutter J, Ulia L;
XX WPI; 2003-147394/14.
XX N-FSDB; ABX14666.
XX Novel ATP-binding cassette transporter-like polypeptides and
PT polynucleotides useful for diagnosing, preventing, treating disorders

PT involving immune, nervous system, thyroid, hypothalamus and impaired
PT transport of lipids.

XX Claim 13; Fig 2; 149pp; English.

XX The invention relates to an isolated murine and human ATP-binding
CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
CC encoded by the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110 or PTA-
CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
CC vectors, host cells, ABCL binding agents, a selective binding agent or
CC its fragment comprising at least one complementarity determining region
CC (CDR) with specificity for ABCL which (produced by immunising an animal
CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
CC fusion polypeptide, a device comprising a membrane suitable for
CC implantation (permeable to the protein and impermeable to materials
CC detrimental to the cells, and cells encapsulated within the membrane)
CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
CC and modulators are useful for the diagnosis and/or treatment of diseases
CC and conditions involving impaired transport of lipids, including
CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
CC involving functional and trophic disturbances of the nervous system such
CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
CC fibrosis, conditions involving multidrug resistance, conditions involving
CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
CC neutropenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies, autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL
XX

SQ Sequence 2146 AA;

Query Match	100.0%;	Score 10896;	DB 6;	Length 2146;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2100;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	PPLSHHCHFPKPLPSAGTVPWVQGLICNVNNTCFPLTPGEPGRSLNFDNLSRL	60		
DB 47	PPLSHHCHFPKPLPSAGTVPWVQGLICNVNNTCFPLTPGEPGRSLNFDNLSRL	106		
QY 61	ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKOSPLPPMLDVAELTSLRT	120		
DB 107	ADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKOSPLPPMLDVAELTSLRT	166		
QY 121	ESLGLALGOAEPPLHSLLEAEADIAOELLALRSVELRALLORPGTSGPLELSEALCS	180		
DB 167	ESLGLALGOAEPPLHSLLEAEADIAOELLALRSVELRALLORPGTSGPLELSEALCS	226		
QY 181	VRGSSVTGPSLNWYEASDLMEVGQSPESALPOSSLPACSELIGALDSHPLSRLWRR	240		
DB 227	VRGSSVTGPSLNWYEASDLMEVGQSPESALPOSSLPACSELIGALDSHPLSRLWRR	286		
QY 241	LKPLILGKLFPAPTPFTRKLMQVNRTEFELTLRLDRVREWMELGPRIPTFMNDSNVA	300		
DB 287	LKPLILGKLFPAPTPFTRKLMQVNRTEFELTLRLDRVREWMELGPRIPTFMNDSNVA	346		
QY 301	MLQRLLOMDGRQPRPGGRDHNEALRSFLDRGSGGYSWODAHADYCHLVGTLGRVTEC	360		
DB 347	MLQRLLOMDGRQPRPGGRDHNEALRSFLDRGSGGYSWODAHADYCHLVGTLGRVTEC	406		
QY 361	LSLDKLEAAPSEALVSRALQLLAEHRFWAGVWFLGPDSDSDTEHPTDLPDGPGRVRIKI	420		
DB 407	LSLDKLEAAPSEALVSRALQLLAEHRFWAGVWFLGPDSDSDTEHPTDLPDGPGRVRIKI	466		
QY 421	RMDIDVTRTNKIRDRFWDPGFAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	480		
DB 467	RMDIDVTRTNKIRDRFWDPGFAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG	526		
QY 481	LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL	540		

DB 527	LYLOQMPYPCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDTRAMGL	586		
QY 541	SRAVLWGLWFLSCIGPFLLSAALLVLVLKLDLIPYSHPGVWFLFLAAFAVATVTSQFLL	600		
DB 587	SRAVLWGLWFLSCIGPFLLSAALLVLVLKLDLIPYSHPGVWFLFLAAFAVATVTSQFLL	646		
QY 601	SAFESRANLAAACGGIAYFSLYPVLCVAVRDRLPAGGRVAASLLSPVAFGFCESLAL	660		
DB 647	SAFESRANLAAACGGIAYFSLYPVLCVAVRDRLPAGGRVAASLLSPVAFGFCESLAL	706		
QY 661	LEEQSGCAOWHNVTGPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPMN	720		
DB 707	LEEQSGCAOWHNVTGPTADVFSLAQVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPMN	766		
QY 721	FPPRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSRSLSEKRRPSPQPALRG	780		
DB 767	FPPRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSRSLSEKRRPSPQPALRG	826		
QY 781	LSLDFYOGHTTAFLGHNGAGKTTTSLTSLGLFPSPSGSAFIIIGHDVRSSMAAIRPHLGVC	840		
DB 827	LSLDFYOGHTTAFLGHNGAGKTTTSLTSLGLFPSPSGSAFIIIGHDVRSSMAAIRPHLGVC	886		
QY 841	POYNVLFDMLTVDHEHWFYGRLLKLSAAVVGPEQDRLLQDVGLVSKOSVOTRHLSGMQR	900		
DB 887	POYNVLFDMLTVDHEHWFYGRLLKLSAAVVGPEQDRLLQDVGLVSKOSVOTRHLSGMQR	946		
QY 901	KLVSIAIAFVGSGSVILDEPTAGVDPASRRGIWELLKRYREGRTLLSTHHLDEAELLGD	960		
DB 947	KLVSIAIAFVGSGSVILDEPTAGVDPASRRGIWELLKRYREGRTLLSTHHLDEAELLGD	1006		
QY 961	RVAVAGGRLLCCCGSPFLFRHLSGYYLTAVKARLPLTTNEKADTMEGSDVTROBKN	1020		
DB 1007	RVAVAGGRLLCCCGSPFLFRHLSGYYLTAVKARLPLTTNEKADTMEGSDVTROBKN	1066		
QY 1021	GSOGSRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGADHGGSFATILFRELDTLAE	1080		
DB 1067	GSOGSRVGTPLLALVQHWVPGARLVEELPHELVLVLPYTGADHGGSFATILFRELDTLAE	1126		
QY 1081	LRLTGYGISTSLSEEIFLKVVEECAADTMEGSCGOHLCTGIAGLDVTURLKMPPOETA	1140		
DB 1127	LRLTGYGISTSLSEEIFLKVVEECAADTMEGSCGOHLCTGIAGLDVTURLKMPPOETA	1196		
QY 1141	LENGEPAAGSAPETDQSGPDVAGRVQWALTROQLQALLKRLFLARRSRGLFAQVLVP	1200		
DB 1187	LENGEPAAGSAPETDQSGPDVAGRVQWALTROQLQALLKRLFLARRSRGLFAQVLVP	1246		
QY 1201	ALFVGLALVFSLIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGPGRARLLEALLQEAGL	1260		
DB 1247	ALFVGLALVFSLIVPPFGHYPALRLSPTMYGAQVSFFSEDAPGPGRARLLEALLQEAGL	1306		
QY 1261	EEPPVQHSRRFSAPEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAE	1320		
DB 1307	EEPPVQHSRRFSAPEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAEVAE	1366		
QY 1321	PPQAVTGSCEVVQNLTKRNLSDFLVKTYPRVQGLTKTKWNEVRYGSGFSLGRDPGL	1380		
DB 1367	PPQAVTGSCEVVQNLTKRNLSDFLVKTYPRVQGLTKTKWNEVRYGSGFSLGRDPGL	1426		
QY 1381	PSGELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLSLDAQDSLKIFWNNKGMHSMVAF	1440		
DB 1427	PSGELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLSLDAQDSLKIFWNNKGMHSMVAF	1486		
QY 1441	VNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM	1500		
DB 1487	VNRASNAIILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLSICVVFAM	1546		
QY 1501	SFVPASFTVLIERVTRAKHLQMGGLSPTLVWGLNFMWDMCNLYLPACIVLILFLAFO	1560		
DB 1547	SFVPASFTVLIERVTRAKHLQMGGLSPTLVWGLNFMWDMCNLYLPACIVLILFLAFO	1606		
QY 1561	QRAYVAPANLPALILLLLXGWSITPLWYPAFFSVFSTAYVVLITCINLIFIGINGSNAT	1620		

Db 1607 QRAYVAPANLPAALLLLLLYGWSITLPMYPASFVSPSTAYVVLTLINLFIGINGSMAT 1666
QY 1621 FVLELFSQKQEVSRILKQVFLIIPFHFCIGRLGLDMVRNQAMADAFERLGDROFQSPUR 1680
Db 1667 FVLELFSQKQEVSRILKQVFLIIPFHFCIGRLGLDMVRNQAMADAFERLGDROFQSPUR 1726
QY 1681 WEVVGKNLLAMVIOGFLFLLTLLQHRSQLLPQPRVSLPLGGEDEDAVARERVVQG 1740
Db 1727 WEVVGKNLLAMVIOGFLFLLTLLQHRSQLLPQPRVSLPLGGEDEDAVARERVVQG 1786
QY 1741 ATQGDVLRNLTKVYRGQRPVAVDRCLGIPPGECFGLGNGAGKSTFRWVTGDTLA 1800
Db 1787 ATQGDVLRNLTKVYRGQRPVAVDRCLGIPPGECFGLGNGAGKSTFRWVTGDTLA 1846
QY 1801 SRGEAVLAGHSVARPSAAHLSMGYCPOSDAIFELLTGREHLELLARLURGVPEAQVOTA 1860
Db 1847 SRGEAVLAGHSVARPSAAHLSMGYCPOSDAIFELLTGREHLELLARLURGVPEAQVOTA 1906
QY 1861 GSGLARLGLSWYADRPAGTYSGNKRKLATALLVDPVAVFLDPTTGMDSARRFLWN 1920
Db 1907 GSGLARLGLSWYADRPAGTYSGNKRKLATALLVDPVAVFLDPTTGMDSARRFLWN 1966
QY 1921 SLLAVVGRGRSMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1967 SLLAVVGRGRSMLTSHSMECEALCSRLAIVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQAAAFAAEFFPGSELREAHGRLRFLPFGRCALARVFGELAVHGAHGVDF 2040
Db 2027 PAARSQAAAFAAEFFPGSELREAHGRLRFLPFGRCALARVFGELAVHGAHGVDF 2086
QY 2041 SVSQTMLSEVFLYFSKQDKDEDTBEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2100
Db 2087 SVSQTMLSEVFLYFSKQDKDEDTBEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2146

RESULT 2
AAU04483
ID AAU04483 standard; protein; 2146 AA.
AC AAU04483;
XX
XX
XX 26-SEP-2001 (first entry)
XX Human PD-ATP-binding cassette (PD-ABC) protein form #1.
XX
XX PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
XX peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
XX cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
XX epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;
XX familial high-density lipoprotein deficiency; fatty liver disease;
XX atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
XX alcoholism; retinal degeneration; hypertension; vascular disease.
XX
XX Homo sapiens.
XX
XX WO200153490-A1.
XX
XX 26-JUL-2001.
XX
XX 23-JAN-2001; 2001WO-US002191.
XX
XX 24-JAN-2000; 2000US-0177889P.
XX 30-JUN-2000; 2000US-0215405P.
XX
XX (WARN) WARNER LAMBERT CO.
XX
XX Johns MA, Tafuri SR, Wang M;
XX
XX WPI; 2001-442259/47.
XX N-PSDB; AAS08706.
XX
XX New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
XX of dyslipidaemia, epilepsy and diseases related to abnormal calcium flux.
XX
XX

XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

Claim 10; Page 54-62; 77pp; English.

The sequence represents human PD-ATP-binding cassette (PD-ABC) protein form 1. PD-ABC maps to chromosome 19p13.3 and is expressed in various tissues including spleen, thymus, peripheral blood leukocytes, bone marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to diagnose and treat cardiovascular disorders, inflammatory disorders, dyslipidaemia, epilepsy, diseases related to abnormal calcium flux, coronary artery disease, Tangier's disease, familial high-density lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease, insulin resistance, obesity, alcoholism, retinal degeneration, hypertension and vascular disease. The sequences are also used in drug screening assays

Sequence 2146 AA;

Query Match 99.9%; Score 10883; DB 4; Length 2146;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2098; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKPDPASAGTVPMLQGLICNVNNTCFPOLTPGEPGRLSNFDSLSVRL 60
Db 47 PPLEHHECHFPNKPDPASAGTVPMLQGLICNVNNTCFPOLTPGEPGRLSNFDSLSVRL 106
QY 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRT 166
QY 121 ESLGALGQAQEPHLSLEAAEDLAQELLARSLVELRALLORPRGTSGPILLEALCS 180
Db 167 ESLGALGQAQEPHLSLEAAEDLAQELLARSLVELRALLORPRGTSGPILLEALCS 226
QY 181 VRGSSSTVGPLNAYEASDLMEYGOEPESALPOSSISPACSELIGALDHPSLRLWR 240
Db 227 VRGSSSTVGPLNAYEASDLMEYGOEPESALPOSSISPACSELIGALDHPSLRLWR 286
QY 241 LKPLILGKLLFAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIFFMNDSSNVA 300
Db 287 LKPLILGKLLFAPDTPFTRKLMQVNRTEELTLRDVREVMELGPRIFFMNDSSNVA 346
QY 301 MLQRLQMDDEGRQRPGRDRHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
Db 347 MLQRLQMDDEGRQRPGRDRHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406
QY 361 LSLDKLEAAPSEALVSRAQLLAEHRFWAGVFLGPEDSSDPTHEPTDLPDGPCHVRIKI 420
Db 407 LSLDKLEAAPSEALVSRAQLLAEHRFWAGVFLGPEDSSDPTHEPTDLPDGPCHVRIKI 466
QY 421 RMDIDVVRTNKRDRFWDPGPAADPLTDLRYVMGGFYLLQDLVERAAVRVLSGANPRAG 480
Db 467 RMDIDVVRTNKRDRFWDPGPAADPLTDLRYVMGGFYLLQDLVERAAVRVLSGANPRAG 526
QY 481 LYLOOMPYPYCVDDVFLRVLSRSLPLFLTLAMIYSVTLTVAVREKTRLDRTNRMGL 540
Db 527 LYLOOMPYPYCVDDVFLRVLSRSLPLFLTLAMIYSVTLTVAVREKTRLDRTNRMGL 586
QY 541 SRAVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPYSHPGVVFLLAFAFATVTSFLL 600
Db 587 SRAVLWLGWFLSCIGPFLLSAALLVVLKLGDIILPYSHPGVVFLLAFAFATVTSFLL 646
QY 601 SAFFSRANLAAACGLAYFSLYPLVLCVNRDLRDPAGGRVAAASLLSPVAFGFCESLAL 660
Db 647 SAFFSRANLAAACGLAYFSLYPLVLCVNRDLRDPAGGRVAAASLLSPVAFGFCESLAL 706
QY 661 LEEQEGGAQWNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 720
Db 707 LEEQEGGAQWNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPN 766
QY 721 FPFRRSWCGRPPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRG 780
Db 767 FPFRRSWCGRPPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRG 826

QY 781 LSLDFYOGHITAFILGHNGAGKTTTSLISGLFPSPGSGSAFTILGHDRVSSMAAIRPHLGVC 840
Db 827 LSLDFYOGHITAFILGHNGAGKTTTSLISGLFPSPGSGSAFTILGHDRVSSMAAIRPHLGVC 886
QY 841 PQYNVLFDMITVDEHWFYGRKGLSAAVVGPEQDRLLQDGLVSKOSVOTRHLSSGGMQR 900
Db 887 PQYNVLFDMITVDEHWFYGRKGLSAAVVGPEQDRLLQDGLVSKOSVOTRHLSSGGMQR 946
QY 901 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDBAELLGD 960
Db 947 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKYREGRTLILSTHHLDBAELLGD 1006
QY 961 RVAVVAGRLCCCSPLFLRHLSGGYYLTIVKARLPLTNEKADTDMESGVDTRQEKKN 1020
Db 1007 RVAVVAGRLCCCSPLFLRHLSGGYYLTIVKARLPLTNEKADTDMESGVDTRQEKKN 1066
QY 1021 GSQSRVGTTPOLLALVOHWVPGARLVBEELPHELVLVLPYTGADHSGPATLFRDLDTRLAE 1080
Db 1067 GSQSRVGTTPOLLALVOHWVPGARLVBEELPHELVLVLPYTGADHSGPATLFRDLDTRLAE 1126
QY 1081 LRLTGYGISDTSLEEIFLKVVVECAADTDMEDSCGQHLCGTGIAGLDVTLLRLKMPQETA 1140
Db 1127 LRLTGYGISDTSLEEIFLKVVVECAADTDMEDSCGQHLCGTGIAGLDVTLLRLKMPQETA 1186
QY 1141 LENGEPAASAPETDGGSPDAVGVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1200
Db 1187 LENGEPAASAPETDGGSPDAVGVQWALTRQOLQALLKRFLLARRSRGLFAQIVLP 1246
QY 1201 ALFVGLALVFLSIVPPGCHYPALRLSPTMYGAQVSFFSEADPGPGRARLLEALLQAGL 1260
Db 1247 ALFVGLALVFLSIVPPGCHYPALRLSPTMYGAQVSFFSEADPGPGRARLLEALLQAGL 1306
QY 1261 EEPVQVHSSHRFSAPEVAEVAKVYLAGNMTPEPSPACQSQPGARRLLPDCPAAAGGP 1320
Db 1307 EEPVQVHSSHRFSAPEVAEVAKVYLAGNMTPEPSPACQSQPGARRLLPDCPAAAGGP 1366
QY 1321 PPOQAVTGSGEVQNLGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGL 1380
Db 1367 PPOQAVTGSGEVQNLGRNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGL 1426
QY 1381 PSQGEGRSVEELWALLSPFGGALDRVLKNTAWAHSLLDAQDSLKTFWNNKGWHSVAF 1440
Db 1427 PSQGEGRSVEELWALLSPFGGALDRVLKNTAWAHSLLDAQDSLKTFWNNKGWHSVAF 1486
QY 1441 VNRSNALRLRAHLPFGPARHAHSITTLNHPNLTKQESAAALMASSVDVLVSTCVVFAM 1500
Db 1487 VNRSNALRLRAHLPFGPARHAHSITTLNHPNLTKQESAAALMASSVDVLVSTCVVFAM 1546
QY 1501 SFVPASFTVLIEBVTTRAKHLQIMGGLSPTLYWGNFLDMCNLYVPACIVVLIFLAQ 1560
Db 1547 SFVPASFTVLIEBVTTRAKHLQIMGGLSPTLYWGNFLDMCNLYVPACIVVLIFLAQ 1606
QY 1561 QRAVAPANLPALELLLLLYGWSITPLMYPASFFSPVSTAYVVLTCINLFIGINGSMAT 1620
Db 1607 QRAVAPANLPALELLLLLYGWSITPLMYPASFFSPVSTAYVVLTCINLFIGINGSMAT 1666
QY 1621 FVLSEFSDQLQEVSRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLR 1680
Db 1667 FVLSEFSDQLQEVSRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLR 1726
QY 1681 WEVVGKLLAMVIOGPIFLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVAREERVVOG 1740
Db 1727 WEVVGKLLAMVIOGPIFLFTLLLOHRSOLLPOPRVRSPLLGEDEEDVAREERVVOG 1786
QY 1741 ATQGDVLVRLNLTKVYRGQMPAVDRCLGTPPGCECFGLLVGNVAGKTSIFRWMTGDTLA 1800
Db 1787 ATQGDVLVRLNLTKVYRGQMPAVDRCLGTPPGCECFGLLVGNVAGKTSIFRWMTGDTLA 1846
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQAQTA 1860
Db 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQAQTA 1906
QY 1861 GSGLARLGLSWYADRPAGTYSGGNKKRLATALAVGDPVAVVFLDEPTTGMDPSARRFLWN 1920

Db 1907 GSGLARLGLSWYADRPAGTYSGGNKKRLATALAVGDPVAVVFLDEPTTGMDPSARRFLWN 1966
QY 1921 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1967 SLLAVVREGSRVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 2026
QY 1981 PAARSQPAALFVAAEFPGSSELREAHGRLRFQLPPGRCALARVFGELAVHGAHGVEDF 2040
Db 2027 PAARSQPAALFVAAEFPGSSELREAHGRLRFQLPPGRCALARVFGELAVHGAHGVEDF 2086
QY 2041 SVSQTMLEEVFLYFSKQDQKDETEQKEAGVGVDPAFGLOHPKRVSQFLDDPSTAEITVL 2100
Db 2087 SVSQTMLEEVFLYFSKQDQKDETEQKEAGVGVDPAFGLOHPKRVSQFLDDPSTAEITVL 2146
RESULT 3
AAO14210
ID AAO14210 standard; protein; 2180 AA.
XX AC AAO14210;
XX AC
DT 03-MAY-2002 (first entry)
XX DE Human transporter and ion channel TRICH-27.
XX DE
XX KW Human; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder;
KW cell proliferative disorder; neuroprotective; nootropic;
KW cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
KW gene therapy.
XX KW
OS Homo sapiens.
XX PN WC200204520-A2.
XX XX
PD 17-JAN-2002.
XX PF 05-JUL-2001; 2001WO-US021448.
XX PR 07-JUL-2000; 2000US-0216547P.
PR 14-JUL-2000; 2000US-0218232P.
PR 21-JUL-2000; 2000US-0220112P.
PR 28-JUL-2000; 2000US-0221839P.
XX (INCY-) INCYTE GENOMICS INC.
XX PA
XX PI Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI Burford N, Greene BD, Sanjanwala MS, Baughn WR, Yao MG, Yang J;
PI Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Wallia NK; Lal P;
PI Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
PI Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;
PI Kearney L, Thangavelu K, Das D, Policky JL;
XX WPI; 2002-205969/26.
DR N-PSDB; AAL44693.
XX XX
XX PT New human transporters and ion channel polypeptides for diagnosing,
PT treating or preventing transport, neurological, muscle, immunological and
XX cell proliferative disorders.
XX PT
PS Claim 1; Page 188-192; 230pp; English.
XX XX
CC The present invention provides the protein and coding sequences of a
CC number of human transporter and ion channel proteins, designated TRICH-1-
CC TRICH-32. The sequences can be used in the treatment of transport,
CC neurological, muscle, immunological and cell proliferative disorders. The
CC present sequence is a protein of the invention
XX XX
SQ Sequence 2180 AA;
Query Match 99.9%; Score 10883; DB 5; Length 2180;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	PPLEHHECHFPNKP	PSAGTVPWMLQGLICNVNNTCFPOLTPGEBPGRLSNFDNLSVSRLL 60
Db	81	PPLEHHECHFPNKP	PSAGTVPWMLQGLICNVNNTCFPOLTPGEBPGRLSNFDNLSVSRLL 140
Qy	61	ADARTVLGGSARHT	LAGLKGKIATLRAARSTAOQOPTKQSPLEPPMLDVALLTSILRT 120
Db	141	ADARTVLGGSARHT	LAGLKGKIATLRAARSTAOQOPTKQSPLEPPMLDVALLTSILRT 200
Qy	121	ESLGLALGOAQEP	LHSLLEAAEDLAQELLALRSVLERALLORPRGTSGLLELSEALCS 180
Db	201	ESLGLALGOAQEP	LHSLLEAAEDLAQELLALRSVLERALLORPRGTSGLLELSEALCS 260
Qy	181	VRGPESTVGPSLNWYEASD	LMELVGOBPESALPDSSLSPACSELIGALDHPRLSLRWR 240
Db	261	VRGPESTVGPSLNWYEASD	LMELVGOBPESALPDSSLSPACSELIGALDHPRLSLRWR 320
Qy	241	LKPLILGKLLPAPDPT	PTTRKLMQVNRTPFELTLLRDVREWEMLGPRITTFMNDSSNVA 300
Db	321	LKPLILGKLLPAPDPT	PTTRKLMQVNRTPFELTLLRDVREWEMLGPRITTFMNDSSNVA 380
Qy	301	MLQRLQWQDEGR	RRPPGGRDHMEALRSFLDPGSGGYSWQDAHADVHLVGLTGRVTEC 360
Db	381	MLQRLQWQDEGR	RRPPGGRDHMEALRSFLDPGSGGYSWQDAHADVHLVGLTGRVTEC 440
Qy	361	LSLDKLEAAPSEAL	VSRAQLLAEHRFWAGVVFILGPEDSSDPTTEHPTDLPDGPCHVRIKI 420
Db	441	LSLDKLEAAPSEAL	VSRAQLLAEHRFWAGVVFILGPEDSSDPTTEHPTDLPDGPCHVRIKI 500
Qy	421	RMDIDVVTNRKIRDR	FWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 480
Db	501	RMDIDVVTNRKIRDR	FWDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRLSGANPRAG 560
Qy	481	LYLQOMPEPCYVDD	VFLVLSRSLPLFLTLAWIYSVTLTVKAVREKETRLDTRMAMGL 540
Db	561	LYLQOMPEPCYVDD	VFLVLSRSLPLFLTLAWIYSVTLTVKAVREKETRLDTRMAMGL 620
Qy	541	SRVILWLGWFLSCL	GPFLLSAALLVLVLKGLDILPYSHPGVVFLLAFAFVATVTSQFLL 600
Db	621	SRVILWLGWFLSCL	GPFLLSAALLVLVLKGLDILPYSHPGVVFLLAFAFVATVTSQFLL 680
Qy	601	SAFFSRANLAAACG	LAYFSLYPYVLCVAMRDLRPAGRVVAASLLSPVAFPGFCESLAL 660
Db	681	SAFFSRANLAAACG	LAYFSLYPYVLCVAMRDLRPAGRVVAASLLSPVAFPGFCESLAL 740
Qy	661	LEQEGEQAQHNVT	RTPADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQGYGIPPPWN 720
Db	741	LEQEGEQAQHNVT	RTPADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQGYGIPPPWN 800
Qy	721	PPFRRSYWCGR	PPKSPAPCTPLDPKVLVEEAPPLGSPGVSVRSLEKRRPGSPQPALRG 780
Db	801	PPFRRSYWCGR	PPKSPAPCTPLDPKVLVEEAPPLGSPGVSVRSLEKRRPGSPQPALRG 860
Qy	781	LSLDIFYQGHITA	FLGHNGAGKTTTLLSILSGLFPSPGSGSAFILGHVDRSSMAAIRPHLGV 840
Db	861	LSLDIFYQGHITA	FLGHNGAGKTTTLLSILSGLFPSPGSGSAFILGHVDRSSMAAIRPHLGV 920
Qy	841	POYNVLFDMLT	VDHEHWFYGRUKLGLSAAVVGPEQDRLLQDVLGYSKQSVQTRHLSGGMQR 900
Db	921	POYNVLFDMLT	VDHEHWFYGRUKLGLSAAVVGPEQDRLLQDVLGYSKQSVQTRHLSGGMQR 980
Qy	901	KL.SVAIAFVGGS	QVWVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHLLDEALLGD 960
Db	981	KL.SVAIAFVGGS	QVWVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHLLDEALLGD 1040
Qy	961	RVAVAGGR	LCCCGSPFLRRHLGSGYYLLTVKARLPLTTNEKADTDMEGSVDTROEKKN 1020
Db	1041	RVAVAGGR	LCCCGSPFLRRHLGSGYYLLTVKARLPLTTNEKADTDMEGSVDTROEKKN 1100
Qy	1021	GSQGSRVGT	PQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFPRELDTRLAE 1080
Db	1101	GSQGSRVGT	PQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSPATLFPRELDTRLAE 1160

RESULT 4
ABP52096
ID ABP52096 standard; protein; 2146 AA.

Qy	1081	LRLTGYGISTSLBIEIFLKVVEECAADTDMEDSGSQHLCTGTGIAGLDVTLRLKMPQETA	1140
Db	1161	LRLTGYGISTSLBIEIFLKVVEECAADTDMEDSGSQHLCTGTGIAGLDVTLRLKMPQETA	1220
Qy	1141	LENPEPAGSAPETDQSGPDVAVRGVQGWALTRQQLQALLKRLFLARRSRRLGPAQIVLP	1200
Db	1221	LENPEPAGSAPETDQSGPDVAVRGVQGWALTRQQLQALLKRLFLARRSRRLGPAQIVLP	1280
Qy	1201	ALFVGLALVFSIIVPPFGHYPALRLSPMYGAQVSFFSEDPGDPGRARLLLEALLQEAEL	1260
Db	1281	ALFVGLALVFSIIVPPFGHYPALRLSPMYGAQVSFFSEDPGDPGRARLLLEALLQEAEL	1340
Qy	1261	EEPPVQHSRHRSFAPEVPAEYAKVLASGNWTPESSPACQSCQSPQCARLLPDCPAAAGP	1320
Db	1341	EEPPVQHSRHRSFAPEVPAEYAKVLASGNWTPESSPACQSCQSPQCARLLPDCPAAAGP	1400
Qy	1321	PPQAVTSGGEVQVNLTORNLSDFLVKTYPRLVROGLTKKWNNEVRVGGFSLGDRDGL	1380
Db	1401	PPQAVTSGGEVQVNLTORNLSDFLVKTYPRLVROGLTKKWNNEVRVGGFSLGDRDGL	1460
Qy	1381	PSGQELGRSVEELWALLSPLPGGALDRVLKXLTAWAHSILDAQDSLKIWFNNKGWSWAF	1440
Db	1461	PSGQELGRSVEELWALLSPLPGGALDRVLKXLTAWAHSILDAQDSLKIWFNNKGWSWAF	1520
Qy	1441	VNRASNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVFAM	1500
Db	1521	VNRASNAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSI CVVFAM	1580
Qy	1501	SFVPASFTVLVIERVTRAKHLQLMGGLSPTLYMGNFLMDMNCNLYLPACIVVILFLAPQ	1560
Db	1581	SFVPASFTVLVIERVTRAKHLQLMGGLSPTLYMGNFLMDMNCNLYLPACIVVILFLAPQ	1640
Qy	1561	QRAYVAPANLALLLLLYGWSITPLMYPASFPESVPSTAVVLTCLNLFINGSMAT	1620
Db	1641	QRAYVAPANLALLLLLYGWSITPLMYPASFPESVPSTAVVLTCLNLFINGSMAT	1700
Qy	1621	FVLELFSQKLEQVSRILKQVFLIPPFCGLGRGLIDMVNRQAMADAFERLGRDQSPQLR	1680
Db	1701	FVLELFSQKLEQVSRILKQVFLIPPFCGLGRGLIDMVNRQAMADAFERLGRDQSPQLR	1760
Qy	1681	WEVVGKILLAMVIOGFLPFLTLLOHRSQLLPQPRVRSPLPGLGEDESDVARERVRVG	1740
Db	1761	WEVVGKILLAMVIOGFLPFLTLLOHRSQLLPQPRVRSPLPGLGEDESDVARERVRVG	1820
Qy	1741	ATQGDVLVRLNLTKVYRGORMPAVDRCLGIPPGCEFCGLGVNGAGKTTSTFRMVTGDTLA	1800
Db	1821	ATQGDVLVRLNLTKVYRGORMPAVDRCLGIPPGCEFCGLGVNGAGKTTSTFRMVTGDTLA	1880
Qy	1801	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVAOTA	1860
Db	1881	SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVAOTA	1940
Qy	1861	CSGLARLGLSNVADRPAGTYSGGNKRKLATLALVDPVAVPLDPTTMDPSARRFLWN	1920
Db	1941	CSGLARLGLSNVADRPAGTYSGGNKRKLATLALVDPVAVPLDPTTMDPSARRFLWN	2000
Qy	1921	SLLAVVREGSRVMLTSHSMECEALCSRLATMVNGRFECLGSPQHLKGRFAAGHTLTURV	1980
Db	2001	SLLAVVREGSRVMLTSHSMECEALCSRLATMVNGRFECLGSPQHLKGRFAAGHTLTURV	2060
Qy	1981	PAARSQAAAAFVAABFPQSELRREAGHGRRLRQLPPGRCALARVFGELAVHGAEGHVEDF	2040
Db	2061	PAARSQAAAAFVAABFPQSELRREAGHGRRLRQLPPGRCALARVFGELAVHGAEGHVEDF	2120
Qy	2041	SVSOTMLBEVFLYFSKQDKDETESEKAGVGVDPAPGLQHPKRVVSFLDDPSTAEVTL	2100
Db	2121	SVSOTMLBEVFLYFSKQDKDETESEKAGVGVDPAPGLQHPKRVVSFLDDPSTAEVTL	2180

```

XX AC ABP52096;
XX DT 10-OCT-2002 (first entry)
XX DE Homo sapiens ABC transporter ABCA7 protein SEQ ID NO:48.
XX KW ATP-binding cassette transporter; ABC transporter; modulation; D loop;
XX KW cancer; bacterial infection; fungal infection; protozoal infection;
XX KW antibacterial; fungicide; protozoacide.
XX OS Homo sapiens.
XX FN EPI217066-A1.
XX PD 26-JUN-2002.
XX PF 21-DEC-2000; 2000EP-00870316.
XX PR 21-DEC-2000; 2000EP-00870316.
XX PA (UYGE-) UNIV GENT.
XX DR WPI; 2002-550404/59.
XX PT Modulating activity of ATP-binding cassette (ABC) transporters by
XX PT influencing dimerization of nucleotide binding domains through use of D
XX PT loop sequence of an ABC transporter, or its antisense peptide or peptide
XX PT mimetic.
XX PS Disclosure; Fig 3; 290pp; English.
XX CC The present invention describes a method (M1) for modulating the activity
XX CC of ATP-binding cassette (ABC) transporters by influencing the
XX CC dimerisation of the nucleotide binding domains comprising using: (a) a
XX CC polypeptide (polyP) consisting of 5-50 amino acids comprising the D loop
XX CC sequence of an ABC transporter (ABP52049 to ABP52091); (b) a polyP
XX CC consisting of the D loop sequence of an ABC transporter; (c) a peptide
XX CC mimetic or antisense peptide of (a) or (b). ABC transporters have
XX CC antibacterial, fungicide and protozoacide activities. (M1) is useful for
XX CC selectively modulating the activity of ABC transporters belonging to the
XX CC group of multidrug transporters/P-glycoproteins. Bacterial, fungal or
XX CC protozoal ABC transporters are involved in the infection of a mammal or
XX CC in the induction of resistance to antibiotics or drugs in a mammal. (M1)
XX CC is useful for preventing, treating or alleviating diseases associated
XX CC with functionality of an ABC transporter. ABP52092 to ABP52140 represent
XX CC ABC transporter proteins given in the exemplification of the present
XX CC invention
XX SQ Sequence 2146 AA;
Query Match 99.9%; Score 10881; DB 5; Length 2146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2098; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 PPLEHHECHFPNKP LPSAGTVPM LQGLICNVNNTCFPQ LTPGEPGR LSNFND SLVSRLL 60
DB 47 PPLEHHECHFPNKP LPSAGTVPM LQGLICNVNNTCFPQ LTPGEPGR LSNFND SLVSRLL 106
QY 61 ADAETVLGGASAHET LAGLGLIATLRAARSTAOPTKOSP LPPMDV AELLTSLRT 120
DB 107 ADAETVLGGASAHET LAGLGLIATLRAARSTAOPTKOSP LPPMDV AELLTSLRT 166
QY 121 ESLGLALQAOEPLHSLLEAAEDLAQELLALRS LVELRALIQRPGTSGP LLELSEALCS 180
DB 167 ESLGLALQAOEPLHSLLEAAEDLAQELLALRS LVELRALIQRPGTSGP LLELSEALCS 226
QY 181 VRGSPSTVGPLSNWYEASDLME LYGQEPESALPDSSISPACSELIGALD SHPLSRLLWR 240
DB 227 VRGSPSTVGPLSNWYEASDLME LYGQEPESALPDSSISPACSELIGALD SHPLSRLLWR 286
QY 241 LKPLILGKLLFAPDTPTRK LMAQVNRTFELTILRDVR EWMGLPRIETFMNDSSNVA 300

```

QY 1381 PSQELGRSVEELWALLSPGALDRVLKNTAWAHSLSDAQDSLKIWNKNGHSMVAF 1440
 DB 1427 PSQELGRSVEELWALLSPGALDRVLKNTAWAHSLSDAQDSLKIWNKNGHSMVAF 1486
 QY 1441 VNRSNAIILRAHLPQGRARHAHSITTLNHPNLTKQLSEALMASSVDVLVSIQVVFAM 1500
 DB 1487 VNRSNAIILRAHLPQGRARHAHSITTLNHPNLTKQLSEALMASSVDVLVSIQVVFAM 1546
 QY 1501 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLGNFMDMNCVLPACIVVLIFLAFO 1560
 DB 1547 SFVPASFTLVLEERVTRAKHLQMGGLSPTLYWLGNFMDMNCVLPACIVVLIFLAFO 1606
 QY 1561 QRAYVAPANLPAALLLLLYGWSITPLMYPASFFSVSTAVVLTICNLFINGSMAT 1620
 DB 1607 QRAYVAPANLPAALLLLLYGWSITPLMYPASFFSVSTAVVLTICNLFINGSMAT 1666
 QY 1621 FVLELPSQOKLEVRILKQVFLIPPHFCIGRGLDMVRNQAMADAFERLGRQFQSPUR 1680
 DB 1667 FVLELPSQOKLEVRILKQVFLIPPHFCIGRGLDMVRNQAMADAFERLGRQFQSPUR 1726
 QY 1681 WEVGKNLLAMVIQGPLFLLFTLLQHRSQLPQPRVSLPLGGEEDVABERERVQOG 1740
 DB 1727 WEVGKNLLAMVIQGPLFLLFTLLQHRSQLPQPRVSLPLGGEEDVABERERVQOG 1786
 QY 1741 ATQGDVLVLRNLTQVYRGQMPAVDRCLGIPPGECFGLLVGNAGKTSFPMVTGDTLA 1800
 DB 1787 ATQGDVLVLRNLTQVYRGQMPAVDRCLGIPPGECFGLLVGNAGKTSFPMVTGDTLA 1846
 QY 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVQTA 1860
 DB 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAQVQTA 1906
 QY 1861 GSGLARLGLSWADRPAGTYSGNKRKLATALVGDPAVVFLEDPPTTGMPSARRFLWN 1920
 DB 1907 GSGLARLGLSWADRPAGTYSGNKRKLATALVGDPAVVFLEDPPTTGMPSARRFLWN 1966
 QY 1921 SLLAVVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTIRV 1980
 DB 1967 SLLAVVREGSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTIRV 2026
 QY 1981 PAARSQAAAFAVAFPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAHGVDF 2040
 DB 2027 PAARSQAAAFAVAFPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAHGVDF 2086
 QY 2041 SVSQTMLEEVLYFSKQKQEDTTEQKEAGVVDPAQLQHPKRVSOFLDDPSTAEVTL 2100
 DB 2087 SVSQTMLEEVLYFSKQKQEDTTEQKEAGVVDPAQLQHPKRVSOFLDDPSTAEVTL 2146

RESULT 5

ABU08466
 ID ABU08466 standard; protein; 2146 AA.
 XX
 AC ABU08466;
 XX
 DT 18-JUN-2003 (first entry)
 XX
 DE Human ABCA-SSN protein.
 XX
 KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
 KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
 KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
 KW immunomodulator; immunosuppressive; antiinflammatory; ABCA-SSN;
 KW antiarteriosclerotic.
 XX
 OS Homo sapiens.
 OS
 PN WO2003010315-A1.
 XX
 PD 06-FEB-2003.
 XX
 PF 24-JUL-2002; 2002WO-JP007487.

XX 25-JUL-2001; 2001JP-00224176.
 PR 06-DEC-2001; 2001JP-00372530.
 XX (KYOW) KYOMA HAKKO KOGYO KK.
 PA (KAZU-) KAZUSA DNA RES INST FOUND.
 XX Ueda K, Nakagawa S, Nagase T;
 XX WPI; 2003-239444/23.
 XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
 PT the immune system, applicable in diagnosis of and screening drugs for
 PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
 XX
 PS Claim 18; Page 164-174; 183pp; Japanese.
 XX
 CC The present invention relates to the isolation of human ATP-binding
 CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
 CC polynucleotide sequences encoding them. The protein is applicable in the
 CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
 CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
 CC It may also be used in a method for screening ABCA-SSN inhibitors. The
 CC present sequence represents human ABCA-SSN protein
 XX
 SQ Sequence 2146 AA;
 Query Match 99.9%; Score 10881; DB 6; Length 2146;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2098; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PPLEHHECHFPKPLPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRISNFNDSLSVRL 60
 DB 47 PPLEHHECHFPKPLPSAGTVPWLQGLICNVNNTCFPQLTGCEEPGRISNFNDSLSVRL 106
 QY 61 ADARTVLGCGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 120
 DB 107 ADARTVLGCGASAHRTLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 166
 QY 121 ESLGALGQAQEPPLHSLLEAAEDLAQELLALRSVLRAQLORPGTSGPPELLSEALCS 180
 DB 167 ESLGALGQAQEPPLHSLLEAAEDLAQELLALRSVLRAQLORPGTSGPPELLSEALCS 226
 QY 181 VRGSSSTVGPLNMYEASDLMEVGOEPESALPOSSLSPACSELIGALDHPSLLLWR 240
 DB 227 VRGSSSTVGPLNMYEASDLMEVGOEPESALPOSSLSPACSELIGALDHPSLLLWR 286
 QY 241 LKPLILGKLLFAPDTPPTTRKLMQVNRTEBELTLRLDYRVEMMLGPRIFTFPMNDSNVA 300
 DB 287 LKPLILGKLLFAPDTPPTTRKLMQVNRTEBELTLRLDYRVEMMLGPRIFTFPMNDSNVA 346
 QY 301 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
 DB 347 MLQRLQMDGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406
 QY 361 LSLDKLEAAPSEAAALVSRALQLLAHREFWAGVFLGPDSDSDPTHEPDPDLGPGHVRKI 420
 DB 407 LSLDKLEAAPSEAAALVSRALQLLAHREFWAGVFLGPDSDSDPTHEPDPDLGPGHVRKI 466
 QY 421 RMDIDVTRTNKIRDRFMDPGFAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 480
 DB 467 RMDIDVTRTNKIRDRFMDPGFAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526
 QY 481 LYLOQMPYPCVDDVFRLVLSRSPLFLTLTAWIYSVTLTVKAVVREKTRLDTRWRMGL 540
 DB 527 LYLOQMPYPCVDDVFRLVLSRSPLFLTLTAWIYSVTLTVKAVVREKTRLDTRWRMGL 586
 QY 541 SRVILWLGWFLSCIGPFLLSAALLVILKLDIILPYSHPGVVFLLAFAFATVQTQSFL 600
 DB 587 SRVILWLGWFLSCIGPFLLSAALLVILKLDIILPYSHPGVVFLLAFAFATVQTQSFL 646
 QY 601 SAFPSTRANLAAACGGLAYFSLYPLVLCVAMRDLRDPAGGRVAASLLSPVAFGFCESLAL 660

Db 647 SAPFSRANLAAACGLAYFSLYLVLCVAMWRDLRPAAGRVAAASLLSPVAFGFCESIAL 706
Qy 661 LEEQEGAGQHNVTTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGQYIPBPWN 720
Db 707 LEEQEGAGQHNVTTRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGQYIPBPWN 766
Qy 721 PPFRRSVMCGPRPKSPAPCPTPLDPKVLVEEAPGSLSPGVSPSLEKRPFGSPOPALRG 780
Db 767 PPFRRSVMCGPRPKSPAPCPTPLDPKVLVEEAPGSLSPGVSPSLEKRPFGSPOPALRG 826
Qy 781 LSLDFYOGHITAFIHNAGAGKTTTILSILSGILFPSPGSAFTILGHDVRSSMAAIPHLGVC 840
Db 827 LSLDFYOGHITAFIHNAGAGKTTTILSILSGILFPSPGSAFTILGHDVRSSMAAIPHLGVC 886
Qy 841 PQYNVLFDMITVDHVMFYGRKLGLSAAVVGPEODRLLQDVLGVSKQSVOTRHLGGQMR 900
Db 887 PQYNVLFDMITVDHVMFYGRKLGLSAAVVGPEODRLLQDVLGVSKQSVOTRHLGGQMR 946
Qy 901 KLSVAIAFVCGSQVVIIDEPAGVDPASRRGIWELLILKYREGRTLILSTHHLDEAELLGD 960
Db 947 KLSVAIAFVCGSQVVIIDEPAGVDPASRRGIWELLILKYREGRTLILSTHHLDEAELLGD 1006
Qy 961 RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROBKKN 1020
Db 1007 RVAVVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROBKKN 1066
Qy 1021 GSQGSRVGTQQLLALVQHVPGARLVEELPHELVLVLPYTGADHSGPATILFRELDTRLAE 1080
Db 1067 GSQGSRVGTQQLLALVQHVPGARLVEELPHELVLVLPYTGADHSGPATILFRELDTRLAE 1126
Qy 1081 LRLTYGIGISLSEIEIFLKVVEECAADTDMEDGSCGHGCTGIAGLDVTLRLKWPQOETA 1140
Db 1127 LRLTYGIGISLSEIEIFLKVVEECAADTDMEDGSCGHGCTGIAGLDVTLRLKWPQOETA 1186
Qy 1141 LENGEPAAGSAPETDQSGPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1200
Db 1187 LENGEPAAGSAPETDQSGPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGFAQIVLP 1246
Qy 1201 ALFVGLALVLSLIVPPFGHYPALRLSPTWYGAQVSFSEADPGDPGRARLLEALLQEAGL 1260
Db 1247 ALFVGLALVLSLIVPPFGHYPALRLSPTWYGAQVSFSEADPGDPGRARLLEALLQEAGL 1306
Qy 1261 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNWTPESPSPACQSQBARLLIPDCPAAGGP 1320
Db 1307 EEPVQVHSSHRFSAPEVAEPAEVAKVLASGNWTPESPSPACQSQBARLLIPDCPAAGGP 1366
Qy 1321 PPOAVTGSGEVQNLGTGRNLSDFLVKTYPRLVQGLTKKVNVEVRYGGSFLGGRDPGL 1380
Db 1367 PPOAVTGSGEVQNLGTGRNLSDFLVKTYPRLVQGLTKKVNVEVRYGGSFLGGRDPGL 1426
Qy 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSDAQDSLKIWFNNKGWHSWVAF 1440
Db 1427 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSDAQDSLKIWFNNKGWHSWVAF 1486
Qy 1441 VNRASNAILRAHLPFGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSVICVVFAM 1500
Db 1487 VNRASNAILRAHLPFGPARHAHSITTLNHPNLTKQELFEAALMASSVDVLVSVICVVFAM 1546
Qy 1501 SFVPASTFLVLIBERVTRAKHLQMGGLSPTLYWLGNFMDMCMNLYLPACIVLILFLAFQ 1560
Db 1547 SFVPASTFLVLIBERVTRAKHLQMGGLSPTLYWLGNFMDMCMNLYLPACIVLILFLAFQ 1606
Qy 1561 QRAYVAPANLPAALLLLLLLXGWSITPLMWYPASFPFSPSTAYVVLITCINLFIGINGSMAT 1620
Db 1607 QRAYVAPANLPAALLLLLLLXGWSITPLMWYPASFPFSPSTAYVVLITCINLFIGINGSMAT 1666
Qy 1621 FVLELFSQKLOEVSRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRDQFQSPLR 1680
Db 1667 FVLELFSQKLOEVSRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGRDQFQSPLR 1726
Qy 1681 WEVVGKLLAMVIOGPFLLFTLLQHRSSOLLPOPRVRSPLLGEEDVVARERVRVQ 1740
Db 1727 WEVVGKLLAMVIOGPFLLFTLLQHRSSOLLPOPRVRSPLLGEEDVVARERVRVQ 1786

RESULT 6

AAU09174

ID AAU09174 standard; protein; 2144 AA.

XX AAU09174;

AC AAU09174;

XX DT DT

XX DT 09-APR-2002 (first entry)

XX DE Human transporter molecule, MTP-1.

XX DE

XX KW

KW Human; transporter molecule; MTP-1; cytotostatic; Nootropic; HIV;

KW Neuroprotective; Antiparkinsonian; Anticonvulsant; Antianaemic; protein;

KW Antidiabetic; Antiartherosclerotic; Anti-human immunodeficiency virus;

KW Antiarthritic; Immunosuppressive; Antitubercular; Tuberculostatic;

KW Antiulcer; Antimanic; Tranquilliser; Vasotropic; fertility disorder;

KW transporter-associated disorder; haematopoietic disorder; anxiety;

KW leukocytic disorder; lipid metabolism; abnormal vascularisation;

KW immunological disorder; inflammatory disease; neurological disorder;

KW obsessive-compulsive disorder; cardiac-related disorder;

KW hormonal disorder; reproductive disorder.

XX KW

OS Homo sapiens.

XX PN

XX WO200187978-A2.

XX XX

XX PD

XX 22-NOV-2001.

XX PF

XX 14-MAY-2001; 2001WO-US015533.

XX XX

XX PR

XX 12-MAY-2000; 2000US-0204211P.

XX XX

XX PA

XX (MILL-) MILLENNIUM PHARM INC.

XX PI

XX Gluckemann M, Curtis RAJ;

XX XX

XX DR

XX WPI; 2002-082985/11.

XX XX

XX N-PSDB; AAS19207.

XX PT

PT New membrane transport protein and polynucleotides, useful for diagnosing

PT and treating transport protein related disorders e.g. cancer, restenosis,

PT asthma and Alzheimer's disease and to identify modulators of therapeutic

PT use.

XX PT

XX PS

PS Claim 12; Fig 1; 141pp; English.

XX CC

CC The invention relates to an isolated membrane transporter protein-1 (MTP-

CC 1) (I). (I) is useful for identifying a compound which modulates the

CC activity of (I). The method comprises contacting (I) or cell expressing
CC (I) with a test compound and determining whether (I) binds to the test
CC compound or determining the effect of the compound on the activity or
CC expression of (I). The identified compound is useful in treatment and
CC diagnosis of a subject having disorders characterised by aberrant or
CC unwanted MTP-1 protein or nucleic acid expression or activity, where
CC transporter-associated disorders include haematopoietic disorders,
CC leukocytic disorders, disorders related to lipid metabolism, disorders
CC involving abnormal vascularisation, immunological disorders, inflammatory
CC diseases, neurological disorders, anxiety disorders, obsessive-compulsive
CC disorders, cardiac-related disorders. Disorders also include cellular
CC proliferation, growth, differentiation, hormonal disorders and
CC reproductive or fertility disorders. The present sequence represents the
CC amino acid sequence of human transporter molecule, MTP-1
XX
SQ Sequence 2144 AA;

Query Match 99.9%; Score 10880; DB 5; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPOLTPGEEPCRLSNFNDSLSRL 60
DB |||||
QY 45 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPOLTPGEEPCRLSNFNDSLSRL 104
DB |||||
QY 61 ADARTVLGASAHRTLAGLGLKIATIRAAARSTAQOPTKQSPLEPPMLDVAELTSLRT 120
DB |||||
QY 105 ADARTVLGASAHRTLAGLGLKIATIRAAARSTAQOPTKQSPLEPPMLDVAELTSLRT 164
DB |||||
QY 121 ESLGALGAQAPPLSHLEAEDLAQELLALRSVLRELALRQPGTSGPIELLSALCS 180
DB |||||
QY 165 ESLGALGAQAPPLSHLEAEDLAQELLALRSVLRELALRQPGTSGPIELLSALCS 224
DB |||||
QY 181 VRGSSVTGPSLNWTEASDLMEVQEPESALPDSSLPACSELGALDHPHLSLLWR 240
DB |||||
QY 225 VRGSSVTGPSLNWTEASDLMEVQEPESALPDSSLPACSELGALDHPHLSLLWR 284
DB |||||
QY 241 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEELTLRDVREVMELGPRIFTFMDSSNVA 300
DB |||||
QY 285 LKPLILGKLLFAPDTPFTFKLMAQVNRNTEELTLRDVREVMELGPRIFTFMDSSNVA 344
DB |||||
QY 301 MLQRLQWQDEGRORPGGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGLGRVTEC 360
DB |||||
QY 345 MLQRLQWQDEGRORPGGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGLGRVTEC 404
DB |||||
QY 361 LSLDKLEAAPSEALVSALQLLAHREHWAGVVFILGPEDSSDPTHEPTDILGPGHVRKI 420
DB |||||
QY 405 LSLDKLEAAPSEALVSALQLLAHREHWAGVVFILGPEDSSDPTHEPTDILGPGHVRKI 464
DB |||||
QY 421 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 480
DB |||||
QY 465 RMDIDVTRTNKIRDRFWDPGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 524
DB |||||
QY 481 LYLOMPYPCYVDDVFLRVLSRPLFLTLAMIYSVTITVKAVREKETRLRDTWRAMGL 540
DB |||||
QY 525 LYLOMPYPCYVDDVFLRVLSRPLFLTLAMIYSVTITVKAVREKETRLRDTWRAMGL 584
DB |||||
QY 541 SRAVLWLGWFLSCLGPFLLSALLVLKLDILPYSHPGVVFLFLAFAVATVTSFLL 600
DB |||||
QY 585 SRAVLWLGWFLSCLGPFLLSALLVLKLDILPYSHPGVVFLFLAFAVATVTSFLL 644
DB |||||
QY 601 SAFFSRANLAAACGLAYFSLYPLVLCVAVNRDLRPGGRVAASLLSPVAFGFCESLAL 660
DB |||||
QY 645 SAFFSRANLAAACGLAYFSLYPLVLCVAVNRDLRPGGRVAASLLSPVAFGFCESLAL 704
DB |||||
QY 661 LEEQEGQAWHNVTPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGOYGIPEPN 720
DB |||||
QY 705 LEEQEGQAWHNVTPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGOYGIPEPN 764
DB |||||
QY 721 FPFRRSYWCGRPPKSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRRFPGSPQPALRG 780
DB |||||
QY 765 FPFRRSYWCGRPPKSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRRFPGSPQPALRG 824
DB |||||

QY 781 LSLDFYQGHITAFILGHNGAGKTTTILSILSGLPPSGSAFILLHDVRSRMAAIRPHLVC 840
DB |||||
QY 825 LSLDFYQGHITAFILGHNGAGKTTTILSILSGLPPSGSAFILLHDVRSRMAAIRPHLVC 884
DB |||||
QY 841 POYNVLFDMLTVDENHWFYGRILKGLSAAVGPEDQRLLDQVGLVSKQSVOTRHLSCGMQ 900
DB |||||
QY 885 POYNVLFDMLTVDENHWFYGRILKGLSAAVGPEDQRLLDQVGLVSKQSVOTRHLSCGMQ 944
DB |||||
QY 901 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGWEILLKYREGRTILSTHLLDEALLGD 960
DB |||||
QY 945 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGWEILLKYREGRTILSTHLLDEALLGD 1004
DB |||||
QY 961 RVAVVAGRLCCCGSPLFLRRHLGSGYYITLVKARLPLTTNEKADTDMEGSVDTROEKN 1020
DB |||||
QY 1005 RVAVVAGRLCCCGSPLFLRRHLGSGYYITLVKARLPLTTNEKADTDMEGSVDTROEKN 1064
DB |||||
QY 1021 GSQSRVGTPLLALVQHWVPGARLVEBELPHELVLVLPYTGADHDSFATLFELEDTRLAE 1080
DB |||||
QY 1065 GSQSRVGTPLLALVQHWVPGARLVEBELPHELVLVLPYTGADHDSFATLFELEDTRLAE 1124
DB |||||
QY 1081 LRLTGYGSDTSLBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1140
DB |||||
QY 1125 LRLTGYGSDTSLBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPQETA 1184
DB |||||
QY 1141 LENGEPAAGSAPETDQSGDPDAVGRVQGWALTRQQLQALLKFLARRSRRLGFAQIVLP 1200
DB |||||
QY 1185 LENGEPAAGSAPETDQSGDPDAVGRVQGWALTRQQLQALLKFLARRSRRLGFAQIVLP 1244
DB |||||
QY 1201 ALFVGLALVFSLIVPPFGHYPALRLSPYMGQAQVSFFSEDAPDGPGRARLLLEALLQEAGL 1260
DB |||||
QY 1245 ALFVGLALVFSLIVPPFGHYPALRLSPYMGQAQVSFFSEDAPDGPGRARLLLEALLQEAGL 1304
DB |||||
QY 1261 EEPVQVSHSRHSPAPAEVPAEVAKVLASGNWTPESPACOCQSPCARLLPDCPAAAGCP 1320
DB |||||
QY 1305 EEPVQVSHSRHSPAPAEVPAEVAKVLASGNWTPESPACOCQSPCARLLPDCPAAAGCP 1364
DB |||||
QY 1321 PPPQAVTSGVGVQNLTCRNLSDFLVKTYPRLVROGLTKKWNVNRVYGGFSLGCRDPGL 1380
DB |||||
QY 1365 PPPQAVTSGVGVQNLTCRNLSDFLVKTYPRLVROGLTKKWNVNRVYGGFSLGCRDPGL 1424
DB |||||
QY 1391 PSGQLGRSVBELWALLSPLPGGALDRVLKNLTAWAHSIDQDSLKIWNKNGHSMWAF 1440
DB |||||
QY 1425 PSGQLGRSVBELWALLSPLPGGALDRVLKNLTAWAHSIDQDSLKIWNKNGHSMWAF 1484
DB |||||
QY 1441 VNRSNATLRAHLPGPBARHAHSITTLNHPNLNLTKEQISEALMASSVDVVLVSIQVVFAM 1500
DB |||||
QY 1485 VNRSNATLRAHLPGPBARHAHSITTLNHPNLNLTKEQISEALMASSVDVVLVSIQVVFAM 1544
DB |||||
QY 1501 SFVPASFTVLIEBRVTRAKHLQMLGGLSPTLYWLGNEFLWDMCNVLPACIWLIFLAFQ 1560
DB |||||
QY 1545 SFVPASFTVLIEBRVTRAKHLQMLGGLSPTLYWLGNEFLWDMCNVLPACIWLIFLAFQ 1604
DB |||||
QY 1561 QRAYVAPANLPAALLLLLYGWSITPLMPYSPFSPFVSTAYVVLTCINLFIGINGSMAT 1620
DB |||||
QY 1605 QRAYVAPANLPAALLLLLYGWSITPLMPYSPFSPFVSTAYVVLTCINLFIGINGSMAT 1664
DB |||||
QY 1621 FVLELFSQDKLOEVSRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAPRELGRQPSPLR 1680
DB |||||
QY 1665 FVLELFSQDKLOEVSRLIKQVFLIFPHFCLGRGLIDMVRNQAMADAPRELGRQPSPLR 1724
DB |||||
QY 1681 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPVRVSLPLGDEEDVARERERVQ 1740
DB |||||
QY 1725 WEVVGKLLAMVIOGPIFLFTLLQHRSQLLPVRVSLPLGDEEDVARERERVQ 1784
DB |||||
QY 1741 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFTFMVTDTLA 1800
DB |||||
QY 1785 ATQGDVLVRLNLTKVYRGORMPAVDRLCLGIPPGECFGLLVNGAGKTSFTFMVTDTLA 1844
DB |||||
QY 1801 SRGAVLGHSVAREPSSAHLSCYCQSDAIPELLTGREHLELLARLGRVPEAQVATA 1860
DB |||||
QY 1845 SRGAVLGHSVAREPSSAHLSCYCQSDAIPELLTGREHLELLARLGRVPEAQVATA 1904
DB |||||
QY 1861 GSGLARLGLSWYADRPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMPDSARRFLWN 1920
DB |||||

Db 1905 GSGLARGLSWADRPAGTSGGNKRLATATALVGDPAVFLDEPTTGNDPSARFLWN 1964
Qy 1921 SLIAVVEGRSVMLTSHMBECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Db 1965 SLLAVVEGRSVMLTSHMBECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 2024
Qy 1981 PAARSQAAAFAVAEFPGSELREAHGRLRFPOLPPGRCALARVFGELAVHGAHGVEDF 2040
Db 2025 PAARSQAAAFAVAEFPGSELREAHGRLRFPOLPPGRCALARVFGELAVHGAHGVEDF 2084
Qy 2041 SVSQTMLIEVFLFSKQDKDETEQKEAGVGVDPAQGLQHPKRVSQFLDDPSTAETVL 2100
Db 2085 SVSQTMLIEVFLFSKQDKDETEQKEAGVGVDPAQGLQHPKRVSQFLDDPSTAETVL 2144

RESULT 7

ADD37429
ID ADD37429 standard; protein; 2144 AA.

XX AC ADD37429;
XX DT 15-JAN-2004 (first entry)
XX DE Human transporter MTP-1.
XX KW Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
XX KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
XX KW diabetes.

XX OS Homo sapiens.
XX PN US2003143675-A1.
XX PD 31-JUL-2003.
XX PF 22-MAY-2002; 2002US-00154419.

XX 12-MAY-2000; 2000US-0204211P.
XX 29-JUN-2000; 2000US-0215376P.
XX 31-JUL-2000; 2000US-0221769P.
XX 19-SEP-2000; 2000US-0233790P.
XX 25-SEP-2000; 2000US-0235107P.
XX 05-OCT-2000; 2000US-0238336P.
XX 14-NOV-2000; 2000US-0248364P.
XX 15-NOV-2000; 2000US-0248878P.
XX 15-DEC-2000; 2000US-0256240P.
XX 18-DEC-2000; 2000US-0256588P.
XX 21-DEC-2000; 2000US-0258028P.
XX 22-JAN-2001; 2001US-0263169P.
XX 14-MAY-2001; 2001US-00858194.
XX 29-JUN-2001; 2001US-00895811.
XX 31-JUL-2001; 2001US-00919781.
XX 19-SEP-2001; 2001US-00957664.
XX 25-SEP-2001; 2001US-00964295.
XX 05-OCT-2001; 2001US-00972724.
XX 14-NOV-2001; 2001US-00002769.
XX 17-DEC-2001; 2001US-00024623.
XX 22-JAN-2002; 2002US-00055025.
XX (MILL-) MILLENNIUM PHARM INC.

XX Curtis RAJ, Glucksmann MA, Meyers RE;
XX WPI; 2003-851783/79.
XX N-PSDB; ADD37428, ADD37430.
XX New isolated nucleic acid, useful for preparing a composition for
XX treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
XX or diabetes.
XX Claim 11; SEQ ID NO 2; 663pp; English.

CC The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment encoding a fragment of the
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.
XX

Sequence 2144 AA;

Query Match 99.9%; Score 10880; DB 7; Length 2144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2097; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PPLEHECHFPNKPSPAGTVPMLQGLICNVNNTCPQLTPGEEPGRLSNFNDLSVRL 60
Db 45 PPLEHECHFPNKPSPAGTVPMLQGLICNVNNTCPQLTPGEEPGRLSNFNDLSVRL 104
Qy 61 ADARTVLGGASARTLAGLKLATLRAARSTAQPTKOSPLEPMLDVAELLTSLRT 120
Db 105 ADARTVLGGASARTLAGLKLATLRAARSTAQPTKOSPLEPMLDVAELLTSLRT 164
Qy 121 ESIGLALGQAQEPHLSILLEAAEDLAQELLALRSIVELRALLQRPRTSGLELLSEALCS 180
Db 165 ESIGLALGQAQEPHLSILLEAAEDLAQELLALRSIVELRALLQRPRTSGLELLSEALCS 224
Qy 181 VRGFSSTVGPSLNWYEASDLMELVQGEPEESALPDSSISPACSELIGALDHPHLSRLWRR 240
Db 225 VRGFSSTVGPSLNWYEASDLMELVQGEPEESALPDSSISPACSELIGALDHPHLSRLWRR 284
Qy 241 LKPLILGLKLLFADPTPTFKLMAQVNTFEELTLDRDVRVEMWMLGPRIETFNWDSNVA 300
Db 285 LKPLILGLKLLFADPTPTFKLMAQVNTFEELTLDRDVRVEMWMLGPRIETFNWDSNVA 344
Qy 301 MLQRLQMQDEGRQRPGRGRDHEALRSFLDPGSGYSWQDAHADVGHVGLGRVTEC 360
Db 345 MLQRLQMQDEGRQRPGRGRDHEALRSFLDPGSGYSWQDAHADVGHVGLGRVTEC 404
Qy 361 LSLDKLEAAPSEAAALYSRALQLLAHRFWAGVVFVGEDSSDTEHPTDLPDGPCHVRIKI 420
Db 405 LSLDKLEAAPSEAAALYSRALQLLAHRFWAGVVFVGEDSSDTEHPTDLPDGPCHVRIKI 464
Qy 421 RMDIDVTRTNKTRDRFWDGPAADPLTDLYVWGGFVYLQDLVERAAVRLSGANPRAG 480
Db 465 RMDIDVTRTNKTRDRFWDGPAADPLTDLYVWGGFVYLQDLVERAAVRLSGANPRAG 524
Qy 481 LXLQOMPYPYCVDDVDFLRLVLSRSLPLFLTLAWYSVTLTKAVVREKETRLDRTMRAMGL 540
Db 525 LXLQOMPYPYCVDDVDFLRLVLSRSLPLFLTLAWYSVTLTKAVVREKETRLDRTMRAMGL 584
Qy 541 SRAVLWLGWFLSLCLGPFLLSAALLVLVLKLDILPYSHPGVVFLLAAFAVATVTSQFL 600
Db 585 SRAVLWLGWFLSLCLGPFLLSAALLVLVLKLDILPYSHPGVVFLLAAFAVATVTSQFL 644
Qy 601 SAFFSRANLAAACGLLAYFSLYLPLPYVLCVAVWRDLRPAAGRVAAASLLSPVAFGCESLAL 660
Db 645 SAFFSRANLAAACGLLAYFSLYLPLPYVLCVAVWRDLRPAAGRVAAASLLSPVAFGCESLAL 704
Qy 661 LEEQGECAQWHNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGQYGPBPWN 720
Db 705 LEEQGECAQWHNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGQYGPBPWN 764
Qy 721 FPFRRSYWCGPRPKSPAPCPTPLDPKVIIVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 780
Db 765 FPFRRSYWCGPRPKSPAPCPTPLDPKVIIVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 824

Qy	761	LSLDPYQGHITAFIGHNGAGKTTTLSILSGLFPSPGGSAGFILGHDVRSMAAIRPHLGV	840
Db	825	:	884
		LSLDPYQGHITAFILGHSAGKRTTTLISILGLFPSPGGSAGFILGHDVRSMAAIRPHLGV	
Qy	841	PQYNVLPDMLTVDEHVWFYGRKGLSAAVGPEDRLLQDVGLVSKQSVOTRHLSGQMOR	900
Db	885	PQYNVLPDMLTVDEHVWFYGRKGLSAAVGPEDRLLQDVGLVSKQSVQTRHLSGQMOR	944
Qy	901	KLVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKKYREGRTLILSTHHLDEAELLGD	960
Db	945	KLVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKKYREGRTLILSTHHLDEAELLGD	1004
Qy	961	RVAVAGRLCCCCSPFLRHLGSGYYLTIVKARLPITTNEKADTDMEGSVDTQBKK	1020
Db	1005	KVAVAGRLCCCCSPFLRHLGSGYYLTIVKARLPITTNEKADTDMEGSVDTQBKK	1064
Qy	1021	GSQSRVGTPLLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRLAE	1080
Db	1065	GSQSRVGTPLLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLFPRELDTRLAE	1124
Qy	1081	LRLTGYGIDSLSLBEIFLKVVEECAADTDMEDGSCQHLCGTIAGLDVTLRLKMPQOETA	1140
Db	1125	LRLTGYGIDSLSLBEIFLKVVEECAADTDMEDGSCQHLCGTIAGLDVTLRLKMPQOETA	1184
Qy	1141	LENCEPAGSAPETDQSGPDAGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP	1200
Db	1185	LENCEPAGSAPETDQSGPDAGVQGWALTRQOLQALLKRFLLARRSRGLFAQIVLP	1244
Qy	1201	ALFVGLALVFSGLIVPPFGHPALRLSPMTYGAQVSFFSEDPAGDPGRARLEALLQEBAGL	1260
Db	1245	ALFVGLALVFSGLIVPPFGHPALRLSPMTYGAQVSFFSEDPAGDPGRARLEALLQEBAGL	1304
Qy	1261	BEPPVQSHSHRFSAPAEVAKVLAISGNWTPESPSPACQSPGARLLLPDCAAAAGP	1320
Db	1305	BEPPVQSHSHRFSAPAEVAKVLAISGNWTPESPSPACQSPGARLLLPDCAAAAGP	1364
Qy	1321	PPQAVTSGSEVQNLTRGNLSDPLVKTYPRLVQGLTKTKWNEVRYGFSLGGRDPGL	1380
Db	1365	PPQAVTSGSEVQNLTRGNLSDPLVKTYPRLVQGLTKTKWNEVRYGFSLGGRDPGL	1424
Qy	1381	PSGQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLDAQDSLTKWFNNKGWHSVAF	1440
Db	1425	PSGQELGRSVEELWALLSPLPGGALDRVLKNLTAWAHSLDAQDSLTKWFNNKGWHSVAF	1484
Qy	1441	VNRASNAILRAHLPPGPAPHAHSITTLNHPILNTKEQLSBAALMASSVDVLVSI	1500
Db	1485	VNRASNAILRAHLPPGPAPHAHSITTLNHPILNTKEQLSBAALMASSVDVLVSI	1544
Qy	1501	SFVPASFTVLVLEERVTRAKHLQWGGSLPTLYWLGNFMDMNCNLYLPACIVLILFLAQ	1560
Db	1545	SFVPASFTVLVLEERVTRAKHLQWGGSLPTLYWLGNFMDMNCNLYLPACIVLILFLAQ	1604
Qy	1561	QRAYVAPANLPALLLLLLLXGWSITPLMWSPASFFSVPESTAYVVLTCINLFIGINGSMAT	1620
Db	1605	QRAYVAPANLPALLLLLLLXGWSITPLMWSPASFFSVPESTAYVVLTCINLFIGINGSMAT	1664
Qy	1621	FVLELFSQDKLOEVSRIILKQVFLIFPHFCLGRGLDMVRNQAMADAPERLGDQFQSPPLR	1680
Db	1665	FVLELFSQDKLOEVSRIILKQVFLIFPHFCLGRGLDMVRNQAMADAPERLGDQFQSPPLR	1724
Qy	1681	WEVVGKLLAMVIOGQPLFTLLLOHRSOLLQPRVRSPLLGEDEDEVARRERVVQ	1740
Db	1725	WEVVGKLLAMVIOGQPLFTLLLOHRSOLLQPRVRSPLLGEDEDEVARRERVVQ	1784
Qy	1741	ATQGDVLVLRNLTKVYRGQRMPAVDRLCLGIPGCECFGLLVGNAGKTSFTFRMVTGDTLA	1800
Db	1785	ATQGDVLVLRNLTKVYRGQRMPAVDRLCLGIPGCECFGLLVGNAGKTSFTFRMVTGDTLA	1844
Qy	1801	SRGAIVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGRHEHLLARLGVPEAQVAQTA	1860
Db	1845	SRGAIVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGRHEHLLARLGVPEAQVAQTA	1904

Qy	1861	GSGLIARIGLSWYADRPAGTYSGGNKRKLATALALVGDPVAVFLDEPTTGMDPSARRFLWN	1920
Db	1905	GSGLIARIGLSWYADRPAGTYSGGNKRKLATALALVGDPVAVFLDEPTTGMDPSARRFLWN	1964
Qy	1921	SLLAVREGSRVMLTSHSMECECALCSRLAIWNGRFRCLGSPHQLKGRFAAGHTLTLRV	1980
Db	1965	SLLAVREGSRVMLTSHSMECECALCSRLAIWNGRFRCLGSPHQLKGRFAAGHTLTLRV	2024
Qy	1981	PAARSQPAARFAVAEPPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAEHGVEDF	2040
Db	2025	PAARSQPAARFAVAEPPGSELREAHGGRRLRFQLPPGRCALARVFGELAVHGAEHGVEDF	2084
Qy	2041	SVSQTMLLEEVFLYPSKQGDDETEQKEAGVGVDPAAGLQHPKRVSQFLDDPSTAEVTVL	2100
Db	2085	SVSQTMLLEEVFLYPSKQGDDETEQKEAGVGVDPAAGLQHPKRVSQFLDDPSTAEVTVL	2144
RESULT 8			
ABUS4629			
ID	ABUS4629	standard; protein; 2059 AA.	
XX	AC	ABUS4629;	
XX	AC	ABUS4629;	
DT	03-JUN-2003	(first entry)	
XX			
XX		Human NOVX polypeptide #88.	
XX			
KW		Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;	
KW		hypertension; congenital heart defect; aortic stenosis; valve disease;	
KW		atrial septal defect; atrioventricular canal defect; ductus arteriosus;	
KW		pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;	
KW		tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;	
KW		obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;	
KW		Parkinson's disease; immune disorder; haematopoietic disorder;	
KW		haemophilia; hypercoagulation; Crohn's disease; cancer.	
OS		Homo sapiens.	
XX			
XX		WO200281498-A2.	
XX			
PD		17-OCT-2002.	
XX			
PF		03-APR-2002; 2002WO-US010780.	
XX			
PR		03-APR-2001; 2001US-0281086P.	
PR		03-APR-2001; 2001US-0281136P.	
PR		05-APR-2001; 2001US-0281863P.	
PR		05-APR-2001; 2001US-0281906P.	
PR		06-APR-2001; 2001US-0282020P.	
PR		10-APR-2001; 2001US-0282930P.	
PR		10-APR-2001; 2001US-0282934P.	
PR		12-APR-2001; 2001US-0283512P.	
PR		13-APR-2001; 2001US-0283710P.	
PR		17-APR-2001; 2001US-0284234P.	
PR		19-APR-2001; 2001US-0285325P.	
PR		20-APR-2001; 2001US-0285381P.	
PR		20-APR-2001; 2001US-0285609P.	
PR		23-APR-2001; 2001US-0285748P.	
PR		23-APR-2001; 2001US-0285890P.	
PR		24-APR-2001; 2001US-0286068P.	
PR		25-APR-2001; 2001US-0286292P.	
PR		27-APR-2001; 2001US-0287213P.	
PR		02-MAY-2001; 2001US-0288257P.	
PR		29-MAY-2001; 2001US-0294164P.	
PR		30-MAY-2001; 2001US-0294484P.	
PR		18-JUN-2001; 2001US-0298952P.	
PR		19-JUN-2001; 2001US-0299237P.	
PR		19-JUN-2001; 2001US-0299276P.	
PR		12-SEP-2001; 2001US-0318750P.	
PR		25-SEP-2001; 2001US-0324800P.	
PR		25-SEP-2001; 2001US-0324802P.	
PR		27-SEP-2001; 2001US-0325684P.	
PR		17-OCT-2001; 2001US-0330143P.	

PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX (CURA-) CURAGEN CORP.
XX
XX Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Patuturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;
PI Gorman L, Shenoy SG, Pena CEA, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shmukets RA, Gangolli EA, Taupier RJ, Casman SU, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
XX WPI; 2003-046858/04.
DR N-PSDB; ABX72257.
XX
XX New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 281-282; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
CC of the invention
XX
SQ Sequence 2059 AA;
Query Match 95.4%; Score 10392.5; DB 6; Length 2059;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 PPLEHHECHFPNKP,PSAGTVPWLOGLICNVNNTCFPQLTGPEPGRLSNFNDSLVRLL 60
Db 47 PPLEHHECHFPNKP,PSAGTVPWLOGLICNVNNTCFPQLTGPEPGRLSNFNDSLVRLL 106
Qy 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 120
Db 107 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRT 166
Qy 121 ESLGALGQAQEP,PHSLLEAAEDLAQELLALRSIVELRALLQRPRGTSGPLELLEALCS 180
Db 167 ESLGALGQAQEP,PHSLLEAAEDLAQELLALRSIVELRALLQRPRGTSGPLELLEALCS 226
Qy 181 VRGPSSSTVGPLNKNYEASDLMEI,VGQBPESALPSSISLSPACSELIGALDHSPLSRLWRR 240
Db 227 VRGPSSSTVGPLNKNYEASDLMEI,VGQBPESALPSSISLSPACSELIGALDHSPLSRLWRR 286
Qy 241 LKPLILGLKLPAPPTPFRKLMAQVNRTELTLLRDVREVMELGPRIFTFMNDSSNVA 300
Db 287 LKPLILGLKLPAPPTPFRKLMAQVNRTELTLLRDVREVMELGPRIFTFMNDSSNVA 346
Qy 301 MLQRLLQMDQEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQADAHVGLVGTIGRVTEC 360
Db 347 MLQRLLQMDQEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQADAHVGLVGTIGRVTEC 406
Qy 361 LSLDKLEAASEAALVSRALQLLAHRFPWAGVFLGPESSDPTTEHPTDLP,GPCHVRIKI 420

Db 407 LSLDKLEAASEAALVSRALQLLAHRFPWAGVFLGPESSDPTTEHPTDLP,GPCHVRIKI 466
Qy 421 RMDIDVVTRTNKIRDRFMDPGPAADPLTDLRYVWGVFVYLODLVERAAVRLSGANPRAG 480
Db 467 RMDIDVVTRTNKIRDRFMDPGPAADPLTDLRYVWGVFVYLODLVERAAVRLSGANPRAG 526
Qy 481 LYLOQMPYPCYVDDVFLVLSRSILPLFLTLAWTYSVTLTVKAUVREKETRLRDTMRWGL 540
Db 527 LYLOQMPYPCYVDDVFLVLSRSILPLFLTLAWTYSVTLTVKAUVREKETRLRDTMRWGL 586
Qy 541 SRAVLWGLWFLSCLGPELISAALLVLVLKGLDILPYSHPGVWFLEAFAFVATVTSQFLL 600
Db 587 SRAVLWGLWFLSCLGPELISAALLVLVLKGLDILPYSHPGVWFLEAFAFVATVTSQFLL 646
Qy 601 SAFFSRANLAAACGLAYFSLYLPYVLCVAWRDLRDPAGGRVAAASLSPVAFGFCESLAL 660
Db 647 SAFFSRANLAAACGLAYFSLYLPYVLCVAWRDLRDPAGGRVAAASLSPVAFGFCESLAL 706
Qy 661 LERQGEQAQWHNVGTRPTADVFSIAQVSGLLLLDAALYGLATWYLEAVCPGOYGIPEPMN 720
Db 707 LERQGEQAQWHNVGTRPTADVFSIAQVSGLLLLDAALYGLATWYLEAVCPGOYGIPEPMN 766
Qy 721 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 780
Db 767 FPFRRSYWCGRPRPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 826
Qy 781 LSLDFYQGHITAFIGHNGAGKTTTSLILGFPSPGSAFILGHVRSMAAIRPHLGVLC 840
Db 827 LSLDFYQGHITAFIGHNGAGKTTTSLILGFPSPGSAFILGHVRSMAAIRPHLGVLC 886
Qy 841 PQYNVLFDMLTVDHWFYGRILKLSNAVVGPDRLLODVLVSKOSVOTRHLSCGMQR 900
Db 887 PQYNVLFDMLTVDHWFYGRILKLSAAVVGPDRLLODVLVSKOSVOTRHLSCGMQR 946
Qy 901 KLSVAIAFVGSGSVIILDEPTAGVDPASRRGIWELLKYREGRTLIILSTHLEDAELLGD 960
Db 947 KLSVAIAFVGSGSVIILDEPTAGVDPASRRGIWELLKYREGRTLIILSTHLEDAELLGD 1006
Qy 961 RVAVAGGRILCCCGSPFLRRHLGSGYLLTVKARPLTTTNEKADTDMESVDTROBKKN 1020
Db 1007 RVAVAGGRILCCCGSPFLRRHLGSGYLLTVKARPLTTTNEKADTDMESVDTROBKKN 1066
Qy 1021 GSQGSRVGTPLLALVQHWPVGARLVEELPHELVLVLPYTGANDGSAFATLRELDTLAE 1080
Db 1067 GSQGSRVGTPLLALVQHWPVGARLVEELPHELVLVLPYTGANDGSAFATLRELDTLAE 1126
Qy 1081 LRLTGYGISDTSLEEIPFKVVECAADTDMEDGSCGOHLCTGTAGLDVTLRLKMPQETA 1140
Db 1127 LRLTGYGISDTSLEEIPFKVVECAADTDMEDGSCGOHLCTGTAGLDVTLRLKMPQETA 1186
Qy 1141 LENGEPAGSAPETDQSGGPDAGVRQGWALTRQOLQALLKRLFLAARRRRLGFAQIVLP 1200
Db 1187 LENGEPAGSAPETDQSGGPDAGVRQGWALTRQOLQALLKRLFLAARRRRLGFAQIVLP 1246
Qy 1201 ALFVGLALVFSILVPPFGHYPALRSETMYGAOVSPESDAPGDPGRARLLLEALQAGL 1260
Db 1247 ALFVGLALVFSILVPPFGHYPALRSETMYGAOVSPESDAPGDPGRARLLLEALQAGL 1306
Qy 1261 BEPPVQHSRHFSAPEVAEPAEVAKVLAGSNWTPSPSPACOCQSGARLLPDCPAAAGGP 1320
Db 1307 BEPPVQHSRHFSAPEVAEPAEVAKVLAGSNWTPSPSPACOCQSGARLLPDCPAAAGGP 1366
Qy 1321 PPQAVTSGSEVQNLGTRNLSDFLVKTYPRLVQRGLTKKWNVEVRYGFSLGGRDPGL 1380
Db 1367 PPQAVTSGSEVQNLGTRNLSDFLVKTYPRLVQRGLTKKWNVEVRYGFSLGGRDPGL 1426
Qy 1381 PSGQELGRSEVELWALLSPLPGGALDRVLKNLTAWAHSLOADSLKTKWNNKWHWSVAF 1440
Db 1427 PSGQELGRSEVELWALLSPLPGGALDRVLKNLTAWAHSLOADSLKTKWNNKWHWSVAF 1486
Qy 1441 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNITKEQLSEAAALMASSVDLVSTCVVFAM 1500

Db 1487 VNRASNAIRLAHLPQPARHAHSITTLNHLNLTKEQLSEAAALMASSVDVLVLSICVVFAM 1546
Qy 1501 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLGFLWDMCNVLPACIVVLFLFAQ 1560
Db 1547 SFVPASFTLVLEERVTRAKHLQLMGGLSPTLYWLGFLWDM----- 1588
Qy 1561 QRAYVAPANLPAALLLLLLYGWSITPLMYPASFFSVPTAYAVLTCTINLFIGINGSMAT 1620
Db 1589 ----- 1588
Qy 1621 FVLELFSDOKLOEVSRIILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQSPRLR 1680
Db 1589 -----KLOEVSRIILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQSPRLR 1639
Qy 1681 WEVVGKLLAMVITQGFLLFTLLLOHRSQLLPQPRVRSPLPGLGEDEDEVARERERVQ 1740
Db 1640 WEVVGKLLAMVITQGFLLFTLLLOHRSQLLPQPRVRSPLPGLGEDEDEVARERERVQ 1699
Qy 1741 ATQGDVLRNLTKVYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSFRMVTGDTILA 1800
Db 1700 ATQGDVLRNLTKVYRGQMPAVDRCLGIPPGECFGLLVNGAGKTSFRMVTGDTILA 1759
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLELLARLRCVPEAQVQTA 1860
Db 1760 SRGEAVLAGHSVAREPSAAHLSMGYCQSDAIFELLTGREHLELLARLRCVPEAQVQTA 1819
Qy 1861 GSGLARLGLSWADRPAGTYSNGKRRKLTALALVGDPAVPLDPTTGMDSARRFLWN 1920
Db 1820 GSGLARLGLSWADRPAGTYSNGKRRKLTALALVGDPAVPLDPTTGMDSARRFLWN 1879
Qy 1921 SLLAVVREGSVMLTSHSMECEALCSRLAIMVNGFRFCGLSPQHLKGRFAAGHTLTLRV 1980
Db 1880 SLLAVVREGSVMLTSHSMECEALCSRLAIMVNGFRFCGLSPQHLKGRFAAGHTLTLRV 1939
Qy 1981 PAARQPAAFVAAFPFSGSELREAGGRRLRFOLPPGRCALARVFGELAVHGAHGVEDF 2040
Db 1940 PAARQPAAFVAAFPFSGSELREAGGRRLRFOLPPGRCALARVFGELAVHGAHGVEDF 1999
Qy 2041 SVSQTMLEEVLYFSKQDKEDTEQKEAGVGVDPAPGLQHPKRVSOPLDDPSTAEVTL 2100
Db 2000 SVSQTMLEEVLYFSKQDKEDTEQKEAGVGVDPAPGLQHPKRVSOPLDDPSTAEVTL 2059

RESULT 9
ABU08464
ID ABU08464 standard; protein; 2008 AA.
XX
AC ABU08464;
XX
DT 18-JUN-2003 (first entry)
XX
DE Amino acid sequence for human ABCA7 splice variant #1.
XX
KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic.
XX
OS Homo sapiens.
XX
FH Key
PT Region
PT 1..28
PT /note= "Given as SEQ ID No:9 and specifically claimed in
PT Claim 2"
XX
WO2003010315-A1.
PN
XX
PD 06-FEB-2003.
XX
PP 24-JUL-2002; 2002WO-JP007487.
XX
PR 25-JUL-2001; 2001JP-00224176.

PR 06-DEC-2001; 2001JP-00372530.
XX
PA (KYOM) KYOMA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
DR N-PSDB; ABX95283.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 1; Page 106-115; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #1
XX
SQ Sequence 2008 AA;
Query Match 94.2%; Score 10264; DB 6; Length 2008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 121 ESLGLALQAOEPLHSLLEAAEDLAQELLALRLSLVELRALLQRPRTSGPLELLSEALCS 180
Db 29 ESLGLALQAOEPLHSLLEAAEDLAQELLALRLSLVELRALLQRPRTSGPLELLSEALCS 88
Qy 181 VRGSSSTVGPSLNVYEAASDLMEVLQGPESALPDSSSLSPACSELTGALDHPRLRLWRR 240
Db 89 VRGSSSTVGPSLNVYEAASDLMEVLQGPESALPDSSSLSPACSELTGALDHPRLRLWRR 148
Qy 241 LKPLTLGLKLLFAPDPTFTRKLMQAVNRTFEELTLRLDRVREVMELGPRIFTPMNDSSNVA 300
Db 149 LKPLTLGLKLLFAPDPTFTRKLMQAVNRTFEELTLRLDRVREVMELGPRIFTPMNDSSNVA 208
Qy 301 MLQRLLOMQDEGRQRPRPGGRDHMEALRSFLDPGSGYSWQDAHADVGLHVLGTLGRVTEC 360
Db 209 MLQRLLOMQDEGRQRPRPGGRDHMEALRSFLDPGSGYSWQDAHADVGLHVLGTLGRVTEC 268
Qy 361 LSLDKLEAAPSEALVSRALQLLAEHRFWAGVVFGLGPDSSDPTPHPTDGLGPGHVRKI 420
Db 269 LSLDKLEAAPSEALVSRALQLLAEHRFWAGVVFGLGPDSSDPTPHPTDGLGPGHVRKI 328
Qy 421 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLODLVERAAVRVLSGANPRAG 388
Qy 481 LYLOQMPYPCVDDVFLRVLSRSPLFLTAWIYSVTILTVKAVVREKETRLDRMTAMGL 540
Db 389 LYLOQMPYPCVDDVFLRVLSRSPLFLTAWIYSVTILTVKAVVREKETRLDRMTAMGL 448
Qy 541 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGDILPYSHPGVVVFLFAAFATVTTQSFL 600
Db 449 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGDILPYSHPGVVVFLFAAFATVTTQSFL 508
Qy 601 SAFTSRANLAACGLAYFSLYLPVLCVNRDLRIPAGRVAASLLSVAFPGCESLAL 660
Db 509 SAFTSRANLAACGLAYFSLYLPVLCVNRDLRIPAGRVAASLLSVAFPGCESLAL 568
Qy 661 LEEQEGEAQWNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGOYGIPEPN 720
Db 569 LEEQEGEAQWNVGTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGOYGIPEPN 628
Qy 721 FPFRRSYWCGPRPKSPACPTPLDPKVLVEAPPGLSPGVSVRSLEKRPFGSPQALRG 780
Db 629 FPFRRSYWCGPRPKSPACPTPLDPKVLVEAPPGLSPGVSVRSLEKRPFGSPQALRG 688

QY 781 LSLDFYOGHITAFILGHNGAGKTTTILSGLFPSPGSGSAFILGHNDVRSSMAAIRPHLGVC 840
Db 689 LSLDFYOGHITAFILGHNGAGKTTTILSGLFPSPGSGSAFILGHNDVRSSMAAIRPHLGVC 748
QY 841 PQYNVLFOMLTVDEHVWFYGRKGLSAAVGPEDQRLLDVGLVSKOSVOTRHLSSGMQR 900
Db 749 PQYNVLFOMLTVDEHVWFYGRKGLSAAVGPEDQRLLDVGLVSKOSVOTRHLSSGMQR 808
QY 901 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDAEALLGD 960
Db 809 KLSVAIAFVGSGQVILDEPTAGVDPASRRGIWELLKLYREGRTLILSTHHLDAEALLGD 868
QY 961 RVAVVAGRLCCCGSPFLRRHLSSGYLTLVKARLPLTNEKADTMEGSDVTRQEKKN 1020
Db 969 RVAVVAGRLCCCGSPFLRRHLSSGYLTLVKARLPLTNEKADTMEGSDVTRQEKKN 928
QY 1021 GSQSRVGTPOLLALVOHVWPGARLVEELPHELVLVLPYTGADHSGSATLFRLELDTLAE 1080
Db 929 GSQSRVGTPOLLALVOHVWPGARLVEELPHELVLVLPYTGADHSGSATLFRLELDTLAE 988
QY 1081 LRLTGYGIDSLSLEIFLKVVVECAADTDMEDSCGQHLCTGTIAGLDVTLRLKMPPOETA 1140
Db 989 LRLTGYGIDSLSLEIFLKVVVECAADTDMEDSCGQHLCTGTIAGLDVTLRLKMPPOETA 1048
QY 1141 LENGEPAAGSAPETDQSGPDVAVGVOGWALTRQLOALLLKRLFLARRSRRLGFAQIVLP 1200
Db 1049 LENGEPAAGSAPETDQSGPDVAVGVOGWALTRQLOALLLKRLFLARRSRRLGFAQIVLP 1108
QY 1201 ALFVGLALVFSLIVPPFGHYPALRLSPMYGAQVSFFSEADPGPGRARLLEALLQAGL 1260
Db 1109 ALFVGLALVFSLIVPPFGHYPALRLSPMYGAQVSFFSEADPGPGRARLLEALLQAGL 1168
QY 1261 BEPPVQSHSHRFAPEVAEPAEVAKLASGNWTPESPACQSPGARRLLPDCPAAAGGP 1320
Db 1169 BEPPVQSHSHRFAPEVAEPAEVAKLASGNWTPESPACQSPGARRLLPDCPAAAGGP 1228
QY 1321 PPQAVTGSGEVQNLGRNLSDFLVKTYPRLVQGLKTKKWNVEVRYGGFSLGGRDPGL 1380
Db 1229 PPQAVTGSGEVQNLGRNLSDFLVKTYPRLVQGLKTKKWNVEVRYGGFSLGGRDPGL 1288
QY 1381 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSAODSLKTFWNNKGWHSWAF 1440
Db 1289 PSGQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLSAODSLKTFWNNKGWHSWAF 1348
QY 1441 VNRAASAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAUMASSVDVLVSVCFWAM 1500
Db 1349 VNRAASAILRAHLPFGPARHAHSITTLNHPNLNLTKEQLSEAAUMASSVDVLVSVCFWAM 1408
QY 1501 SFVPASFTVLIIIEERVTRAKHLQMGGLSPTLYWLGNFMDMCMYIVPACIVLIFLAQ 1560
Db 1409 SFVPASFTVLIIIEERVTRAKHLQMGGLSPTLYWLGNFMDMCMYIVPACIVLIFLAQ 1468
QY 1561 QRAVAPANLPAALLLLLLYGSITPLMYPASFFSPSTAYVVLTCINLFIGINGSWAT 1620
Db 1469 QRAVAPANLPAALLLLLLYGSITPLMYPASFFSPSTAYVVLTCINLFIGINGSWAT 1528
QY 1621 FVLELFSQKLQEVSRILKQVFLIFPHFLGRGLDMVRNQAMADAFERLGDQFQSPLR 1680
Db 1529 FVLELFSQKLQEVSRILKQVFLIFPHFLGRGLDMVRNQAMADAFERLGDQFQSPLR 1588
QY 1681 WEVVGKLLAMVIOGPIFLPFTLLQHRSQLLPQPRVRSPLIGEEEDVARERERVQ 1740
Db 1589 WEVVGKLLAMVIOGPIFLPFTLLQHRSQLLPQPRVRSPLIGEEEDVARERERVQ 1648
QY 1741 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPPCECFGLLGVNGAGKSTFRWVTGDTLA 1800
Db 1649 ATQGDVLVRLNLTKVYRGORMPADVRLCLGTPPCECFGLLGVNGAGKSTFRWVTGDTLA 1708
QY 1801 SRGEAVLAGHSVAREPSAAHLSMGVCFQSDAIFELLTGREHLELLARLGRVPEAQVTA 1860
Db 1709 SRGEAVLAGHSVAREPSAAHLSMGVCFQSDAIFELLTGREHLELLARLGRVPEAQVTA 1768

QY 1861 GSGLARLGLSWADRPAGTYSGGNKRKLATALAVGDPVAVVFLDEPTTGMDDPSARRFLWN 1920
Db 1769 GSGLARLGLSWADRPAGTYSGGNKRKLATALAVGDPVAVVFLDEPTTGMDDPSARRFLWN 1828
QY 1921 SLLAVVREGSRVMTLSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1980
Db 1829 SLLAVVREGSRVMTLSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRV 1888
QY 1981 PAARSQPAAFVAAEPFGSELREAHGRLRFQLPPOGRCALARVFGELAVHGAHGVEDF 2040
Db 1889 PAARSQPAAFVAAEPFGSELREAHGRLRFQLPPOGRCALARVFGELAVHGAHGVEDF 1948
QY 2041 SVSQTMLVEEFLYFSKQDQKDETEOEKAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2100
Db 1949 SVSQTMLVEEFLYFSKQDQKDETEOEKAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2008

RESULT 10
ABU08465
ID ABU08465 standard; protein; 1993 AA.
XX
AC ABU08465;
XX
DT 18-JUN-2003 (first entry)
XX
DE Amino acid sequence for human ABCA7 splice variant #2.
XX
KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
KW immunomodulator; immunosuppressive; antiinflammatory;
KW antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO2003010315-A1.
XX
PD 06-FEB-2003.
XX
PF 24-JUL-2002; 2002WO-JP007487.
XX
PR 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX
DR WPI; 2003-239444/23.
DR N-PSDB; ABX95284.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 1; Page 145-154; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #2
XX
SQ Sequence 1993 AA;

Query Match 93.4%; Score 10171.5; DB 6; Length 1993;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

QY 121 ESLGLALGQAQEPHLSLEAEADLAQELLALRLSIVELRALQLQRPGTSGPSELLSEALCS 180

Db 29 ESIGLALQAOPLSHSLLEAAEDLAQELALSLVELRALQRPRTSGPLELSEALCS 88
Qy 181 VRGPESTVGPNSLWYEAEDLMELVQEPESALPDSLSGPACSELIGALDSHPLSRLWRR 240
Db 89 VRGPESTVGPNSLWYEAEDLMELVQEPESALPDSLSGPACSELIGALDSHPLSRLWRR 148
Qy 241 LKPLILGKLLFAPDPTFTTKLMAQVNRRTFEELTLDRVREWEMLGPRIFTFMDSSNVA 300
Db 149 LKPLILGKLLFAPDPTFTTKLMAQVNRRTFEELTLDRVREWEMLGPRIFTFMDSSNVA 208
Qy 301 MLORLLQODGRRPRGGRDHMALRSPLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 360
Db 209 MLORLLQODGRRPRGGRDHMALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 268
Qy 361 LSLDKLEAAPSEAAALVSALQALLAEHRFWAGVVFGLGPDSSDPTHEPTPDLPBGHVRKI 420
Db 269 LSLDKLEAAPSEAAALVSALQALLAEHRFWAGVVFGLGPDSSDPTHEPTPDLPBGHVRKI 328
Qy 421 RMDIDVTRTKIRDPFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 480
Db 329 RMDIDVTRTKIRDPFWDGPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAG 388
Qy 481 LYLOQMPYPCYVDDVFLRVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRTRAMGL 540
Db 389 LYLOQMPYPCYVDDVFLRVLSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRTRAMGL 448
Qy 541 SRAVLMLGWFLSCGLGPPPLLSAALLVLKIGDILPYSHPGVVFPLAFAFATVTTQSPLL 600
Db 449 SRAVLMLGWFLSCGLGPPPLLSAALLVLKIGDILPYSHPGVVFPLAFAFATVTTQSPLL 508
Qy 601 SAFFSRANLAAACGGGLAYFSLYLPVLCVAMWRDLRPPAGGRVAASLLSPVAFGFCESLAL 660
Db 509 SAFFSRANLAAACGGGLAYFSLYLPVLCVAMWRDLRPPAGGRVAASLLSPVAFGFCESLAL 568
Qy 661 LEEQEGAGQWNVGTRPTADVPFSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYGIPBPWN 720
Db 569 LEEQEGAGQWNVGTRPTADVPFSLAQVSGLLLLDAAALYGLATWYLEAVCPQGYGIPBPWN 628
Qy 721 PPFRRSYWCGPRPPKSPACPTPLDPKVLVEAPPLGPGVSVRSLEKRFPGSPQALRG 780
Db 629 PPFRRSYWCGPRPPKSPACPTPLDPKVLVEAPPLGPGVSVRSLEKRFPGSPQALRG 688
Qy 781 LSLDFYQGHITAFLGHNAGAKTTTILSILSGFPPSPGSAFILGHGHDVRSMAAIRPHLVC 840
Db 689 LSLDFYQGHITAFLGHNAGAKTTTILSILSGFPPSPGSAFILGHGHDVRSMAAIRPHLVC 748
Qy 841 PQYNVLFDMLTVDHFWYFGRKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 900
Db 749 PQYNVLFDMLTVDHFWYFGRKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMQR 808
Qy 901 KLSVAIAFVGGSOVVLDEPTAGVDPASRRGIWELLKLYRSGRTLLILSTHHLDEALIGD 960
Db 809 KLSVAIAFVGGSOVVLDEPTAGVDPASRRGIWELLKLYRSGRTLLILSTHHLDEALIGD 868
Qy 961 RVAVVAGRLCCCGSPLFLRRHLGSGYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN 1020
Db 869 RVAVVAGRLCCCGSPLFLRRHLGSGYTLTVKARLPLTTNEKADTMEGSDVTRQEKKN 928
Qy 1021 GSQSRVCTPOLLALVQHVWPGARLVEBELPHELVLVLPYTGADHGSFATLFEILDTRLAE 1080
Db 929 GSQSRVCTPOLLALVQHVWPGARLVEBELPHELVLVLPYTGADHGSFATLFEILDTRLAE 988
Qy 1081 LRLTCYGISDTSLEBIFLKVVEECAADTDMEDGSCGHLCTGAGLDVTLRLKMPQETA 1140
Db 989 LRLTCYGISDTSLEBIFLKVVEECAADTDMEDGSCGHLCTGAGLDVTLRLKMPQETA 1048
Qy 1141 LENGEPCASAPETDGGSPDAGVRVQGWALTRQOQLALLKRFLLARSRRLGFAQIVLP 1200
Db 1049 LENGEPCASAPETDGGSPDAGVRVQGWALTRQOQLALLKRFLLARSRRLGFAQIVLP 1108
Qy 1201 ALFVGLALVSLIIVPPFGHPALRLSPMTYGAQVSVFFSEADPGDPGRARLLLEALLQEAGL 1260

Db 1109 ALPVLGALVSLIIVPPFGHPALRLSPMTYGAQVSVFFSEADPGDPGRARLLLEALLQEAGL 1168
Qy 1261 EEPVPVQHSRHPESAPEVAEVAKVULASGNWTPESPSPACQSQPGARRLLPCCPAAAGP 1320
Db 1169 EEPVPVQHSRHPESAPEVAEVAKVULASGNWTPESPSPACQSQPGARRLLPCCPAAAGP 1228
Qy 1321 PPPQAVTSGSVEVQVNLTCGRNLSDFLVKTYPRLVRLQGLTKKWKVNVVRVYGGFSLGGRDPL 1380
Db 1229 PPPQAVTSGSVEVQVNLTCGRNLSDFLVKTYPRLVRLQGLTKKWKVNVVRVYGGFSLGGRDPL 1288
Qy 1381 PSQELGHSVEBELWALLSPLPGGALDRVLKNTAWAHSIDAOQSLKIFPNKNGHSMWAF 1440
Db 1289 PSQELGHSVEBELWALLSPLPGGALDRVLKNTAWAHSIDAOQSLKIFPNKNGHSMWAF 1348
Qy 1441 VNRSASNLRAHLPPGPARHAHSITTLNHLNLTKEQLSEAAALMASSVDVLVSI CVVPAM 1500
Db 1349 VNRSASNLRAHLPPGPARHAHSITTLNHLNLTKEQLSEAAALMASSVDVLVSI CVVPAM 1408
Qy 1501 SFVPASPTLVIIEERVTRAKHLQLMGGLSPTLYMGNFLMDMCMNYLVPACIVVLIFLAPQ 1560
Db 1409 SFVPASPTLVIIEERVTRAKHLQLMGGLSPTLYMGNFLMDMCMNYLVPACIVVLIFLAPQ 1468
Qy 1561 QRAYVAPANLPAALLLLLYGNSITPLMYPASFPSPSTAYVLTCLNLFIGINGSMAT 1620
Db 1469 QRAYVAPANLPAALLLLLYGNSITPLMYPASFPSPSTAYVLTCLNLFIGINGSMAT 1528
Qy 1621 FVLELPSQKLOEVSRIKQVPLIFPHFCLGRGLIDMVNRQAMADAPERLGRDQSPULR 1680
Db 1529 FVLELPSQKLOEVSRIKQVPLIFPHFCLGRGLIDMVNRQAMADAPERLGRDQSPULR 1588
Qy 1681 WEVVGKNLLAMVIOGPLFLLTLLQHRSQLLPQVRSLPLLGEEDEDAVARERERVVQG 1740
Db 1589 WEVVGKNLLAMVIOGPLFLLTLLQHRSQLLPQVRSLPLLGEEDEDAVARERERVVQG 1648
Qy 1741 ATQGDVLVRLNLTKVYRGQRMPAVDRCLGIPPGCEFCGLLVNGAGKTTSTPRMVTGDTLA 1800
Db 1649 ATQGDVLVRLNLTKVYRGQRMPAVDRCLGIPPGCEFCGLLVNGAGKTTSTPRMVTGDTLA 1708
Qy 1801 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAOVAQTA 1860
Db 1709 SRGEAVLAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGVPEAOVAQTA 1768
Qy 1861 GSGLARLGLSWYADRPACTYSGGNKRLATALALVGDPAVVFLEPTTGMDSARRFLWN 1920
Db 1769 GSGLARLGLSWYADRPACTYSGGNKRLATALALVGDPAVVFLEPTTGMDSARRFLWN 1828
Qy 1921 SLLAVREGRSVMLTSHSMEECEALCSRLAIWNGRFRCLGSPQHLKGRPAAGHTLTLRV 1980
Db 1829 SLLAVREGRSVMLTSHSMEECEALCSRLAIWNGRFRCLGSPQHLKGRPAAGHTLTLRV 1876
Qy 1981 PAARSQAAAAFVAAEFPGSELREAHGGRRLRFQLPGGRCALARVFGELAVHGAEGHVEDF 2040
Db 1877 ---RSQPAANAFAAEFPQSELREAHGGRRLRFQLPGGRCALARVFGELAVHGAEGHVEDF 1933
Qy 2041 SVSQTMLEVEFLYFSKQDKDEDETEQKEAGVGDPAFGLQHPKRVSVQFLDDPSTAEVTL 2100
Db 1934 SVSQTMLEVEFLYFSKQDKDEDETEQKEAGVGDPAFGLQHPKRVSVQFLDDPSTAEVTL 1993

RESULT 11
AAU04484
ID AAU04484 standard; protein; 1873 AA.
XX
AC AAU04484;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human PD-ATP-binding cassette (PD-ABC) protein form #2.
KW PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;

family high-density lipoprotein deficiency; fatty liver disease; atherosclerosis; diabetes; insulin resistance; obesity; drug screening; alcoholism; retinal degeneration; hypertension; vascular disease.

Homo sapiens.

WO200153490-A1.

26-JUL-2001.

23-JAN-2001; 2001WO-US002191.

24-JAN-2000; 2000US-0177889P.

30-JUN-2000; 2000US-0215405P.

(WARN) WARNER LAMBERT CO.

Johns MA, Tafuri SR, Wang M;

WPI; 2001-442259/47.

N-PSDB; AAS08707.

New Human PD-ABC DNA molecules and proteins for diagnosis and treatment of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.

Claim 10; Page 64-72; 77pp; English.

The sequence represents human PD-ATP-binding cassette (PD-ABC) protein form 2. PD-ABC maps to chromosome 19p13.3 and is expressed in various tissues including spleen, thymus, peripheral blood leukocytes, bone marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to diagnose and treat cardiovascular disorders, inflammatory disorders, dyslipidemia, epilepsy, diseases related to abnormal calcium flux, coronary artery disease, Tangier's disease, familial high-density lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease, insulin resistance, obesity, alcoholism, retinal degeneration, hypertension and vascular disease. The sequences are also used in drug screening assays

Sequence 1873 AA;

Query Match 84.6%; Score 9213.5; DB 4; Length 1873;

Best Local Similarity 97.4%; Pred. No. 0;

Matches 1789; Conservative 6; Mismatches 27; Indels 15; Gaps 4;

Qy 1 PPLEHHECHFPNKPPLPSAGTVPMLQGLICNVNNTCFPQLTPGEPGRLSNFDLSRL 60

Db 47 PPLEHHECHFPNKPPLPSAGTVPMLQGLICNVNNTCFPQLTPGEPGRLSNFDLSRL 106

Qy 61 ADARTVLGGASAHRTLAGLKIATLRAASTAQOPTKOSPPEPMLDVALLTSLRT 120

Db 107 ADARTVLGGASAHRTLAGLKIATLRAASTAQOPTKOSPPEPMLDVALLTSLRT 166

Qy 121 ESLGLALCOQEPPLHSLEAEADLAQELLARSVELLALQRPRTSGPGLLELSEALCS 180

Db 167 ESLGLALCOQEPPLHSLEAEADLAQELLARSVELLALQRPRTSGPGLLELSEALCS 226

Qy 181 VRGFSSTVGPLSNWYEASDLMEVLGQEPESALPDSSLSACSELIGALDHPHLSRLWRR 240

Db 227 VRGFSSTVGPLSNWYEASDLMEVLGQEPESALPDSSLSACSELIGALDHPHLSRLWRR 286

Qy 241 LKPLILGKLIAPDPTTRKLMQVNRFTFELTLRDVREVWEMLGPRIITFMNDSNVA 300

Db 287 LKPLILGKLIAPDPTTRKLMQVNRFTFELTLRDVREVWEMLGPRIITFMNDSNVA 346

Qy 301 MLQRLLQMDQEGRRPPGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360

Db 347 MLQRLLQMDQEGRRPPGGRDHEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 406

Qy 361 LSLDKLEAAPSEALVSRAALQLLAHRFWAGVFLGPEDSSDPTHEPTDPLGFGHVRKI 420

Db 407 LSLDKLEAAPSEALVSRAALQLLAHRFWAGVFLGPEDSSDPTHEPTDPLGFGHVRKI 466

Qy 421 RMDIDVVTRNKIRDRFWDGPAADPLTDLRYVMGFFVYLQDLVERAAVAVLSEANPRAG 480

Db 467 RMDIDVVTRNKIRDRFWDGPAADPLTDLRYVMGFFVYLQDLVERAAVAVLSEANPRAG 526

Qy 481 LYLQMPYPYCYVDDVFLRLVLSRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRMAGL 540

Db 527 LYLQMPYPYCYVDDVFLRLVLSRSLPLFLTLAWIYSVTLTAKAVVREKETRLDTRMAGL 586

Qy 541 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVFLFLAFAVAVTQSFLL 600

Db 587 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVFLFLAFAVAVTQSFLL 646

Qy 601 SAFFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLRPAAGRAASLLSPVAFGFCESLAL 660

Db 647 SAFFSRANLAAACGGLAYFSLYLPYVLCVAVWRDLRPAAGRAASLLSPVAFGFCESLAL 706

Qy 661 LEEQGEQAQWHNVGTRPTADVFSLAQVSGLLLDLAALYGLATWYLEAVCPGOYGIPEPWN 720

Db 707 LEEQGEQAQWHNVGTRPTADVFSLAQVSGLLLDLAALYGLATWYLEAVCPGOYGIPEPWN 766

Qy 721 FPPRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 780

Db 767 FPPRRSYWCGRPPKSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRG 826

Qy 781 LSLDFYQGHITAFIIGHNGAGKTTTLSTLSGLFPPSGSAFIIIGHDVRSSMAAIRPHLGVC 840

Db 827 LSLDFYQGHITAFIIGHNGAGKTTTLSTLSGLFPPSGSAFIIIGHDVRSSMAAIRPHLGVC 886

Qy 841 PQYNVLFDMLTVDHEWVFYGRKLGLSAAVVGPPQDRLLQDVLGVLSQSVQTRHLSGMQR 900

Db 887 PQYNVLFDMLTVDHEWVFYGRKLGLSAAVVGPPQDRLLQDVLGVLSQSVQTRHLSGMQR 946

Qy 901 KLSVAIAFVGGSVIILDEPTAGVDPASRRGIWELLKRYEGRTLLSTHHLDEAELLGD 960

Db 947 KLSVAIAFVGGSVIILDEPTAGVDPASRRGIWELLKRYEGRTLLSTHHLDEAELLGD 1006

Qy 961 RVAVAVAGRLCCCGSPFLFRRHIGSGYLTLVKARLPLTTNEKADTDMEGSVDTROSKN 1020

Db 1007 RVAVAVAGRLCCCGSPFLFRRHIGSGYLTLVKARLPLTTNEKADTDMEGSVDTROSKN 1066

Qy 1021 GSQGSRVGTPQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSGFATLRELDTRLAE 1080

Db 1067 GSQGSRVGTPQLLALVQHWVPGARLVEELPHELVLVLPYTGADGSGFATLRELDTRLAE 1126

Qy 1081 LRLTGYSIDTSLEELFLKVVRECAADTDMEDSCGHLCTIAGLDVTLRLKMPPOETA 1140

Db 1127 LRLTGYSIDTSLEELFLKVVRECAADTDMEDSCGHLCTIAGLDVTLRLKMPPOETA 1186

Qy 1141 LENGEPAAGAPETDQSGPDVAGRVQGWALTROQLQALLKRLFLARRSRRLGFAQIVLP 1200

Db 1187 LENGEPAAGAPETDQSGPDVAGRVQGWALTROQLQALLKRLFLARRSRRLGFAQIVLP 1246

Qy 1201 ALFVGLALVFSLLVPPFGHYPALRLSPTMYGAOVSFSEADPGDGRARLLEALQEAGL 1260

Db 1247 ALFVGLALVFSLLVPPFGHYPALRLSPTMYGAOVSFSEADPGDGRARLLEALQEAGL 1306

Qy 1261 EEPVQVHSHRFSAPVPAEVAKVLASGNWTPSPSPACQSPGARRLLPDCPAAAGGP 1320

Db 1307 EEPVQVHSHRFSAPVPAEVAKVLASGNWTPSPSPACQSPGARRLLPDCPAAAGGP 1366

Qy 1321 PPDQAVTGSVEVQNLTGRLNSDFLVKTYPRVQGLKTKKWNVRYGGFSLGGRDPGL 1380

Db 1367 PPDQAVTGSVEVQNLTGRLNSDFLVKTYPRVQGLKTKKWNVRYGGFSLGGRDPGL 1426

Qy 1381 PSQOELGRSVEELWALLSPLPGGALDRVLKNTAWAHSDAQDSLKIWFNNKGMHSMVAF 1440

Db 1427 PSQOELGRSVEELWALLSPLPGGALDRVLKNTAWAHSDAQDSLKIWFNNKGMHSMVAF 1486

Qy 1441 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSCVVFAM 1500

Db 1487 VNRASNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSCVVFAM 1546

Qy 1501 SFVPAFTLVLIERVTRAKHLQLMGGLSPTLVMLGNFLWDMCNLYLPACIVVLIPLAFQ 1560

Db 1547 SFVPASFTLVLEERVTRAKHLQLMGSLSPITLYWLGNFMDMNVLPACIVLVFLAFQ 1606
 QY 1561 QRAYAPANLPALLLLLLLYGHSITPLMYPASFFSVSTAVVLTINLFTGINGSMAT 1620
 Db 1607 QRAYAPANLPALLLLLLLYGHSITPLMYPASFFSVSTAVVLTINLFTGINGSMAT 1666
 QY 1621 FVLELFSDOKQEVSRILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1680
 Db 1667 FVLELFSDOKQEVSRILKQVFLIPPHFCLGRGLDMVRNQAMADAFERLGRQFSPLR 1726
 QY 1681 MEVWGNLAWYIQOPLFLFTLLQHRSQLPQPRVSLPLGLGEEDVARERERVQO 1740
 Db 1727 MEVWGNLAWYIQOPLFLFTLLQHRSQLPQPRVSLPLGLGEEDVARERERVQO 1786
 QY 1741 ATQGDVLVRLNLTQVYRGQMPAVDRCLGIPGCEGGLGVNGAGKSTFRVNTGDTLA 1800
 Db 1787 ATQGDVLVRLNLTQVYRGQMPAVDRCLGIPGCEVSP--GYEARCRDSEWLPYCMPCPS 1844
 QY 1801 SRGEAVLAGHSVAREPSAAHLSMG-YCQSDAIFELL 1836
 Db 1845 SFTEHLLCIH-----HLLGTYCM---PIFVLL 1869

RESULT 12

ID ABG72695

XX ABG72695 standard; protein; 2167 AA.

AC ABG72695;

XX 10-MAR-2003 (first entry)

DT Mouse ATP-binding cassette transporter-like protein, ABCL.

DE Mouse; ATP-binding cassette transporter-like protein; ABCL;

XX lipid transport; cardiovascular disease; hypertriglyceridaemia;

KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;

KW nervous system disorder; Stargardt disease; degenerative disorder;

KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;

KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;

KW acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;

KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;

KW hypochalumus disorder; obesity; diabetes; reproductive disorder;

KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;

KW autoimmune disease; inflammatory disease; multiple sclerosis.

XX Mus musculus.

OS

XX Key

XX Location/Qualifiers

XX Peptide

XX 1. 46

XX /label= Signal_peptide

XX Protein

XX 47. 2167

XX /label= Mature_ABL

XX US2002127647-A1.

XX 12-SEP-2002.

XX 28-NOV-2001; 2001US-00995542.

XX 28-NOV-2000; 2000US-0253520P.

XX (SHUT/) SHUTTER J.

XX (ULIA/) ULIAS L.

XX Shutter J, Ulia L;

XX WPI; 2003-147394/14.

XX N-PSDB; ABX14665.

XX Novel ATP-binding cassette transporter-like polypeptides and

XX polynucleotides useful for diagnosing, preventing, treating disorders

XX involving immune, nervous system, thyroid, hypothalamus and impaired

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT transport of lipids.

XX Claim 13; Fig 1; 149pp; English.

XX The invention relates to an isolated murine and human ATP-binding cassette transporter-like (ABCL) polypeptide, or the amino acid sequence encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-3111. Also included are the nucleic acids encoding the ABCL proteins, vectors, host cells, ABC binding agents, a selective binding agent or its fragment comprising at least one complementarity determining region (CDR) with specificity for ABCL which (produced by immunising an animal with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL fusion polypeptide, a device comprising a membrane suitable for, an ABCL implantation (permeable to the protein and impermeable to materials detrimental to the cells, and cells encapsulated within the membrane) where the cells secrete ABCL, an ABCL transgenic non-human mammal and an array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids and modulators are useful for the diagnosis and/or treatment of diseases and conditions involving impaired transport of lipids, including CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis, CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions CC involving functional and trophic disturbances of the nervous system such as Stargardt disease, degenerative and inflammatory retinopathy, cystic CC fibrosis, conditions involving multidrug resistance, conditions involving CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias, CC neutropaenia, anaemia and autoimmune diseases, conditions involving the CC thyroid e.g. hyper and hypothyroidism; conditions involving the CC hypothalamus including obesity, diabetes, reproductive disorders, energy CC balance disorders, peripheral neuropathies including myelinopathies and CC axonopathies, autoimmune and inflammatory diseases involving the nervous CC system including multiple sclerosis. The present sequence represents CC murine ABCL

XX Sequence 2167 AA;

Query Match 77.0%; Score 8389.5; DB 6; Length 2167;
 Best Local Similarity 76.3%; Pred. No. 0;
 Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;
 QY 1 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNTCFPQLTGEEPRGLSNFNDSLSRL 60
 Db 47 PPLEHHECHFPNKLPSAGTVPWLQGLICNVNNSCFQHTPGKPGVLSNFKDSLISRL 106
 QY 61 ADARTVLGCGAHRITLGLKLIATLRAARTAQP-----OPTKQSPLEPPMLDVALLTS 116
 Db 107 ADTRTVLGGHSHIQMDLDALGKLI PVLRAVGGARQESDQTSQG-----SVTKLLEK 159
 QY 117 LLRTESLGLAQOQEPHLSLEAAEDLAQELLALRLSLVELRALLRPRGTSGPLELISE 176
 Db 160 ILQASLDPLVGLQADSMRKFSDAIRDLAQELLTLPMLERALLRRPRGSGSLVSE 219
 QY 177 ALCSVRGSSSTVGPSSLNWYEASDLMLVGPBPESALPDSSLSPACSELIGALDSHPLSRL 236
 Db 220 ALCSVKGPSSPGGLSLNWYEANQLNEFMGPEVAPALPDNSLSPACSEPVGTLDHVPVSR 279
 QY 237 LWRRLKPLILGKLLFAPDTPTRKLMQVNTFFELTLRDVREWEMLGRIFTFMDS 296
 Db 280 LWRRLKPLILGKLLFAPDTPTRKLMQVNTFFELTLRDVREWEMLGRIFTFMDS 339
 QY 297 SNVAMLRQLLOQDEGRPRGGRDHMEALRSFLDPGSGGVSWODAHADVGLVGTGR 356
 Db 340 TNVAMLRQLLOQDEGRPRGGRDHMEALRSFLDPGSGGVSWODAHADVGLVGTGR 399
 QY 357 VTECLSLDKLEAAPSEALVSRLQLLAHRFWAGVWVFLGPEDESDPTTEHPTDLPQGHV 416
 Db 400 MMECVSLDKLEAVPSEALVSRLQLLAHRFWAGVWVFLGPEDESDPTTEHPTDLPQGHV 459
 QY 417 RIKRMDIDVTRTNKIRDRFWDGPAADPLTDLYVWGGVFFYLODLVERAARVLSGAN 476
 Db 460 RFIKRMIDIDVTRTNKIRDRFWDGPAADPLTDLYVWGGVFFYLODLVERAARVLSGAN 519
 QY 477 PRAGLYLQQMPYPCVVDVFLVLSRSIPLFLTLTAWYSVTLTKAVVREKTRLRDTR 536

Db 520 SRTGLYLQMPHPCYDDVFLRLVLSRLPLFLTLAWIYVSVALTVKAVVREKETRELTMR 579
Qy 537 AMGLSRAVLNLCWFLSCGLPFLSAAALIVLVKGLDILPYSHPGVWFLFLAAFAVATVQ 596
Db 580 AMGLSRAVLNLCWFLSCGLPFLSAAALIVLVKGLNGLPYSHPVVIFLFLAAFAVATVAQ 639
Qy 597 SFLLSAFPSFRANLAAACGLAYFSLPYLVLCVAVMRDLRPAAGRVAAASLLSPVAFGCGE 656
Db 640 SFLLSAFPSFRANLAAACGLAYFALYLYPYLVLCVAVMRDLRHLGGLAAASLLSPVAFGCGE 699
Qy 657 SLALLEEQGEAQMHNVTGTRPTADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGQYGP 716
Db 700 SLALLEEQGEAQMHNLTGTPAEDVFSLAQVSAFLLDVAVIYGLAWYLEAVCPGQYGP 759
Qy 717 EPNWPPRRSRYWCPCRPKSPAPCTPLDPKVLVSEAPPGLSPGVSRSLSEKFPSPQP 776
Db 760 EPNWPPRRSRYWCPCRPKSPVLPAPQDPKVLVEEPPLGLVPGVSTRGLKGFRCGPQP 819
Qy 777 ALRGLSLDFYQGHITAFIHNAGAKTTTSLTSLGLPPSPGSAFIFLGHVRSMAAIRPH 836
Db 820 ALQGLNLDFYEGHITAFIHNAGAKTTTSLTSLGLPPSPGSAFIFLGHVRSMAAIRPH 879
Qy 837 LGVCPQNVLFDMLTVDVSHVWFYGRKLSAAVVPQDRLLDQVGLVSKQSVQTRHLSG 896
Db 880 LGICPQNVLFDMLTVEHVWFYGRKLSAAVVPQDRLLDQVGLVSKQSVQTRHLSG 939
Qy 897 GMQRKLSVAIAFVGGSOVILDEPTAGVDPASRRGIWELLKLYREGTTLILSTHLLDEAE 956
Db 940 GMQRKLSVAIAFVGGSRVIMDEPTAGVDPASRRGIWELLKLYREGTTLILSTHLLDEAE 999
Qy 957 LLGDRVAVAGRLCCCGSLPLFRHHLGSGVYLTAKARPLTNE-KADTMEGSDVTR 1015
Db 1000 LLGDRVAVAGRLCCCGSLPLFRHHLGSGVYLTAKARPLTNE-KADTMEGSDVTR 1054
Qy 1016 QEKNGSGQS-----RVGTQALLALVQHWVPGARLVEELP 1050
Db 1055 REKSDCNGRTSDTAFRTGTSKSNQAPACAVPITPSTARILLELVQHVPGALVEDLP 1114
Qy 1051 HELVLVLPYTGADHGSFATLPRELDTRLAEURLTGYIGISDTSLEBIFLKVVECAADTM 1110
Db 1115 HELVLVLPYTGADHGSFATLPRELDTRLAEURLTGYIGISDTSLEBIFLKVVECAADTM 1172
Qy 1111 EDGSCGHLCTGIAGLDVTLRLKMPQETALENGEPAGSA--PETDQSGPDVAGRVQGV 1168
Db 1173 GDSRPQLHLRT-----CTQPPTGPESVLENGELAKVLDPQAPQGLAPNA-AQVQGV 1225
Qy 1169 ALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1228
Db 1226 TLTCQQLRALHLKFLARRSRRLGFAQIVLPALFVGLALVFSLIVPPFGHYPALRLSPT 1285
Qy 1229 MYGAQVFFSDEADPGDGRARLLBALLOEAGLEBPVQHSR-----FSAPEV 1277
Db 1286 MYGPQVFFSDEADPGDGRARLLBALLOEAGLEBPVQHSR-----FSAPEV 1345
Qy 1278 PAEVAKVLASGNWTPESPSPACQSQSGARLLPDCPAAAGPPPPQAVTGSERVQNL 1337
Db 1346 PPDVASILASGNWTPESPSPACQSQSGARLLPDCPAAAGPPPPQAVTGSERVQNL 1405
Qy 1338 GRNLSDFLVKTYPLRVLQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELWALL 1397
Db 1406 GRNLSDFLVKTYPLRVLQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELWALL 1465
Qy 1398 SPLPGGALDVLKNTLAWHSLDAQDSLKTWFKNGHSMVAFVNRASNAILRAHLP 1457
Db 1466 SPOGNAIDRLNLTQWALGLDARNSLKTFWFKNGHSMVAFVNRANGLHALLPSGP 1525
Qy 1458 ARHAHSLTTLNHPNLNLTKEQLSEALWASSVDVLVSVVFMASFVPSFTLVLEERT 1517
Db 1526 VRAHSLTTLNHPNLNLTKEQLSEATLJASSVDVLVSVVFMASFVPSFTLVLEERT 1585
Qy 1518 RAKHLQMLGSLPTLYWGLNPLWDMCNLYVPACTIVLILFLAQORAYVAPANLALL 1577
Db 1586 RAKHLQMLGSLPTLYWGLNPLWDMCNLYVAVCIIVLFLAQORAYVAPANLALL 1645

Qy 1578 LLYCWSITPLMYPASRFFSVSTAYVVLTCINLFIGINGSMATFVLELFDOKLOEVSRI 1637
Db 1646 LLYCWSITPLMYPASRFFSVSTAYVVLTCINLFIGINGSMATFVLELFDOKLOEVSRI 1705
Qy 1638 LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVVGNLLAMVIOG 1697
Db 1706 LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVVGNLLAMVIOG 1765
Qy 1698 FLITLTLQHRSQLLPQPRVRSPLGEDEEDVARERVRVQATQGDVILVRLNLTQYR 1757
Db 1766 FLITLTLQHRSQLLPQPRVRSPLGEDEEDVARERVRVQATQGDVILVRLNLTQYR 1825
Qy 1758 GORPAPVDRCLGIPCECEGGLGVNAGCTSTFRMTVGTDLASRGAVALAGHSVAREPS 1817
Db 1826 GORPAPVDRCLGIPCECEGGLGVNAGCTSTFRMTVGTDLASRGAVALAGHSVAREPS 1885
Qy 1818 AAHLSMGYCPQSDAIFELITGRBHELELLARLGRVPEAQAQTAGSGLARLGLSWYADP 1877
Db 1886 AAHLSMGYCPQSDAIFELITGRBHELELLARLGRVPEAQAQTAGSGLARLGLSWYADP 1945
Qy 1878 GTYSGGNKRLATALVGDPAVVFDEPTTGMDDPSARRFLMNSLLAVVREGSRVMTSH 1937
Db 1946 GTYSGGNKRLATALVGDPAVVFDEPTTGMDDPSARRFLMNSLLAVVREGSRVMTSH 2005
Qy 1938 SMECEALCSRLLAMVNGFRCLGSOHLKGRFAAGHTLTLRVPAAASQAAAFVAAEFP 1997
Db 2006 SMECEALCSRLLAMVNGFRCLGSOHLKGRFAAGHTLTLRVPAAASQAAAFVAAEFP 2065
Qy 1998 GSELRAHGRRLFPQLPPGRCALARVFGELAVHGAHGVDEFSVSTMLVEEFLYFSKD 2057
Db 2066 GSELRAHGRRLFPQLPPGRCALARVFGELAVHGAHGVDEFSVSTMLVEEFLYFSKD 2125
Qy 2058 QGKDEDETEQKAGVVDPAQGLQPKRVSQFLDDPSTAEVL 2100
Db 2126 QGKDEDETEQKAGVVDPAQGLQPKRVSQFLDDPSTAEVL 2167
RESULT 13
ABG72697
ID ABG72697 standard; protein; 1550 AA.
XX
AC ABG72697;
XX
XX 10-MAR-2003 (first entry)
XX
XX Human ATP-binding cassette transporter-like protein, ABCCL1550.
DE
XX
XX Human; ATP-binding cassette transporter-like protein; ABCCL1550;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypochloraemia disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
OS
XX Homo sapiens.
XX
XX US2002127647-A1.
PN
XX
PD 12-SEP-2002.
XX
XX 28-NOV-2001; 2001US-00995542.
PF
XX
XX 28-NOV-2000; 2000US-0253520P.
PR
XX
XX (SHUT/) SHUTTER J.
PA (ULIA/) ULIA L.
PA
XX

PI Shutter J, Ulias L;
 DR WPI; 2003-147394/14.
 DR N-PSDB; ABX14667.
 DX
 PT Novel ATP-binding cassette transporter-like polypeptides and
 PT polynucleotides useful for diagnosing, preventing, treating disorders
 PT involving immune, nervous system, thyroid, hypothalamus and impaired
 PT transport of lipids.
 XX
 PS Claim 13; Fig 3; 149pp; English.
 XX
 CC The invention relates to an isolated murine and human ATP-binding
 CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
 CC encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
 CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
 CC vectors, host cells, ABCL binding agents, a selective binding agent or
 CC its fragment comprising at least one complementary determining region
 CC (CDR) with specificity for ABCL which (produced by immunising an animal
 CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
 CC fusion polypeptide, a device comprising a membrane suitable for
 CC implantation (permeable to the protein and impermeable to materials
 CC detrimental to the cells, and cells encapsulated within the membrane)
 CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
 CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
 CC and modulators are useful for the diagnosis and/or treatment of diseases
 CC and conditions involving impaired transport of lipids, including
 CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
 CC hypercholesterolaemia, tangier disease, dyslipidaemias, conditions
 CC involving functional and trophic disturbances of the nervous system such
 CC as Stargard disease, degenerative and inflammatory retinopathy, cystic
 CC fibrosis, conditions involving multidrug resistance, conditions involving
 CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
 CC neutropenia, anaemia and autoimmune diseases, conditions involving the
 CC thyroid e.g. hyper and hypothyroidism; conditions involving the
 CC hypothalamus including obesity, diabetes, reproductive disorders, energy
 CC balance disorders, peripheral neuropathies including myelinopathies and
 CC axonopathies, autoimmune and inflammatory diseases involving the nervous
 CC system including multiple sclerosis. The present sequence represents
 CC human ABCL truncated variant, ABCL1550
 XX
 SX Sequence 1550 AA;
 Query Match 69.7%; Score 7596; DB 6; Length 1550;
 Best Local Similarity 99.6%; Pred No. 0;
 Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 121 ESLGALGQAQPHLSLEAEADLAQLALSLVELRALQPRGTSGPLELLSEALCS 180
 DB 29 ESLGALGQAQPHLSLEAEADLAQLALSLVELRALQPRGTSGPLELLSEALCS 88
 QY 181 VRGSPSTVGPSLNWTEASDLMEVQEPESALPDSLSLSPACSELIGALDHPHLLWRR 240
 DB 89 VRGSPSTVGPSLNWTEASDLMEVQEPESALPDSLSLSPACSELIGALDHPHLLWRR 148
 QY 241 LKPLILGKLLPAPDTPFTRKMAQVNRNFEELTLRDVREVWEMLGPRIFTFMNDSSNVA 300
 DB 149 LKPLILGKLLPAPDTPFTRKMAQVNRNFEELTLRDVREVWEMLGPRIFTFMNDSSNVA 208
 QY 301 MLQRLQMQDEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 360
 DB 209 MLQRLQMQDEGRRQPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTEC 268
 QY 361 LSLDKLEAPSEALVSRAQLLAHREFWAGVVFGLPEDSSDPTHTPDLPGRHVRKI 420
 DB 269 LSLDKLEAPSEALVSRAQLLAHREFWAGVVFGLPEDSSDPTHTPDLPGRHVRKI 328
 QY 421 RMDIDVTRTKIRDFWDGPAADPLTDLRVVWGSGFYVLODLVERAAVRVLSGANPRAG 480
 DB 329 RMDIDVTRTKIRDFWDGPAADPLTDLRVVWGSGFYVLODLVERAAVRVLSGANPRAG 388
 QY 481 LYLOQMPYPCYVDDVFLVLSRLPLFLTLAWIYSVTLTVKA VVREKETRLDTRAMGL 540
 DB

DB 389 LYLOQMPYPCYVDDVFLVLSRLPLFLTLAWIYSVTLTVKA VVREKETRLDTRAMGL 448
 QY 541 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVVFLEAFAVATVTSFLL 600
 DB 449 SRAVLWLGWFLSCIGPFLLSAALLVLVLKGLDILPYSHPGVVFLEAFAVATVTSFLL 508
 QY 601 SAFFSRANLAAACGGGLAYFSLYLPVYLCVAMWDRUPAGGRVAASLLSPVAFGGCSLAL 660
 DB 509 SAFFSRANLAAACGGGLAYFSLYLPVYLCVAMWDRUPAGGRVAASLLSPVAFGGCSLAL 568
 QY 661 LEEQGEQAQHNHVGTRPTADVFSLAQVSGLLLLDAALYGLATWLYLSEAVCPGYGIEPEMN 720
 DB 569 LEEQGEQAQHNHVGTRPTADVFSLAQVSGLLLLDAALYGLATWLYLSEAVCPGYGIEPEMN 628
 QY 721 PFRRSYWCGRPPPKSPAPCPPTLPDPKVLVEAPPGLSPGVSRVSRLEKPPGSPQALRG 780
 DB 629 PFRRSYWCGRPPPKSPAPCPPTLPDPKVLVEAPPGLSPGVSRVSRLEKPPGSPQALRG 688
 QY 781 LSLDFVQGHITAFPHNGAGKTTTLLSGLPFPSPGSAFILLGHVDRSSMAAIRPHLGV 840
 DB 689 LSLDFVQGHITAFPHNGAGKTTTLLSGLPFPSPGSAFILLGHVDRSSMAAIRPHLGV 748
 QY 841 POYNVLFDMLTVDHFWFYGRUKLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGMQR 900
 DB 749 POYNVLFDMLTVDHFWFYGRUKLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGMQR 808
 QY 901 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLLSTHLLDEALLGD 960
 DB 809 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTLLSTHLLDEALLGD 868
 QY 961 RVAVVAGRLCCCGPFLRRHLGSGYTLTKARLPLTTNEKADTDMGSSVDTRKK 1020
 DB 869 RVAVVAGRLCCCGPFLRRHLGSGYTLTKARLPLTTNEKADTDMGSSVDTRKK 928
 QY 1021 SGQSRVGTQQLLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAE 1080
 DB 929 SGQSRVGTQQLLALVQHWPGARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAE 988
 QY 1081 LRLTCYGISDTSLEIFLKVVEECAADTDMGSCGQHLCTGICAGLDVTLRLKMPQETA 1140
 DB 989 LRLTCYGISDTSLEIFLKVVEECAADTDMGSCGQHLCTGICAGLDVTLRLKMPQETA 1048
 QY 1141 LENGEPAQSAPETDQSGPDAVGRVQGWALTRQQLQALLKRLARRRRLGFAQIVLP 1200
 DB 1049 LENGEPAQSAPETDQSGPDAVGRVQGWALTRQQLQALLKRLARRRRLGFAQIVLP 1108
 QY 1201 ALFVGLALVFSLIVPPFGHYPALRLSPMTYGAQVFFSEDAFGDPGRARLLEALQEA 1260
 DB 1109 ALFVGLALVFSLIVPPFGHYPALRLSPMTYGAQVFFSEDAFGDPGRARLLEALQEA 1168
 QY 1261 EEPVQVHSHRPSADEVPAEVAKVLAGSNWTPESPACQSQPGARRLLPDCPAAAGP 1320
 DB 1169 EEPVQVHSHRPSADEVPAEVAKVLAGSNWTPESPACQSQPGARRLLPDCPAAAGP 1228
 QY 1321 PPPQAVTSGGEVQNLGRLNLDLVKTYPRLVROGLTKKWNVVRVYSGSLGGRDGL 1380
 DB 1229 PPPQAVTSGGEVQNLGRLNLDLVKTYPRLVROGLTKKWNVVRVYSGSLGGRDGL 1288
 QY 1381 PSGELGRSVEELWALLSPLFGALDRVLKNTLAWHSLDAQDSLKIWFNNKWSHVA 1440
 DB 1289 PSGELGRSVEELWALLSPLFGALDRVLKNTLAWHSLDAQDSLKIWFNNKWSHVA 1348
 QY 1441 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI 1500
 DB 1349 VNRSNAILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSI 1408
 QY 1501 SFVPASFTLVLIERVTRAKHLQLMGGSLPTLYLGNFLWDMCNLYLVPACI 1560
 DB 1409 SFVPASFTLVLIERVTRAKHLQLMGGSLPTLYLGNFLWDMCNLYLVPACI 1468
 QY 1561 QRAYVAPANLPALELILLYGWSITPL 1587
 DB 1469 QRAYVAPANLPALELILLYGWSITPL 1495

RESULT 14

AAB38111
 ID AAB38111 standard; protein; 2261 AA.
 AC AAB38111;
 CC
 XX 29-JAN-2001 (first entry)
 DD
 DE Human ABC1 cholesterol transporter mutant, V771M.
 XX
 XX Human ABC1 cholesterol transporter; chromosome 9q31;
 KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
 KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
 KW cardiovascular disease; coronary artery disease; coronary restenosis;
 KW cerebrovascular disease; peripheral vascular disease;
 KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
 KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
 KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
 KW mutin.
 XX
 OS Homo sapiens.
 XX
 XX W0200055318-A2.
 XX
 XX 21-SEP-2000.
 XX
 XX 15-MAR-2000; 2000WO-IB000532.
 XX
 XX 15-MAR-1999; 99US-0124702P.
 PR
 PR 08-JUN-1999; 99US-0138048P.
 PR
 PR 17-JUN-1999; 99US-0139600P.
 PR
 PR 01-SEP-1999; 99US-0151977P.
 XX
 XX (UYBR-) UNIV BRITISH COLUMBIA.
 PA
 PA (XENO-) XENON BIORSEARCH INC.
 XX
 XX Hayden MR, Wilson AR, Pimstone SN;
 XX
 XX WPI, 2000-587528/55.
 XX
 XX New ABC1 polypeptide is useful for treating diseases associated with ABC1
 PT biological activity, e.g. Alzheimer's disease, Huntington's disease and
 FT cancer.
 PT
 PT
 XX
 XX Example; Page; 229pp; English.
 XX
 XX The invention relates to the human ABC1 cholesterol transporter protein
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of
 CC proteins, and plays a crucial role in cholesterol transport, particularly
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
 CC located on chromosome 9q31, and mutations in this gene are associated
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
 CC are distinguishable in that TD is an autosomal recessive disorder, while
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
 CC cholesterol") in the blood correlate with a high risk of cardiovascular
 CC disease, particularly coronary artery disease, but also cerebrovascular
 CC disease, coronary restenosis, and peripheral vascular disease.
 CC Conversely, a high level of HDL has protective effects against
 CC cardiovascular disease. The invention provides genetic constructs and
 CC transgenic cells and non-human animals comprising human ABC1 nucleic
 CC acids, and methods of gene therapy for the treatment or prevention of
 CC cardiovascular disease comprising the administration of an expression
 CC vector encoding ABC1 or an active fragment thereof. The invention also
 CC encompasses compounds which mimic ABC1 activity, compounds which
 CC stimulate ABC1 expression and methods of screening for such compounds. It
 CC further relates to methods for determining whether a patient has an
 CC increased risk for cardiovascular disease due to polymorphisms in the
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or

CC prevent cardiovascular disease, especially coronary artery disease,
 CC cerebrovascular disease, coronary restenosis or peripheral vascular
 CC disease. They may also be used in the treatment of diseases associated
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
 CC The invention specifically excludes proteins with the exact amino acid
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
 CC present sequence represents a mutant human ABC1 cholesterol transporter
 CC associated with an altered cholesterol level and therefore an altered
 CC risk of cardiovascular disease. Note: The present sequence is not shown
 CC in the specification, but is derived from the native human ABC1 shown on
 CC pages 152-157
 XX
 SQ Sequence 2261 AA;
 Query Match 51.5%; Score 5611.5; DB 3; Length 2261;
 Best Local Similarity 50.0%; Pred. No. 0;
 Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
 1 PPLEHCHFPNKPPLPSAGTVPWLOGLICNVNTCTPOLTGPGEPLSNFNDLSYRL 60
 47 PPYEQHECHFPNKPPLPSAGTVPWLOGLICNVNTCTPOLTGPGEPLSNFNDLSYRL 106
 61 ADARTVLGGASAHRTLAGLGLKIATLRAARSTAQ-----PQ 96
 107 SDARRLLLYSKDTSMDKMKVLRITLQIKSSNLKLDLVNDETFSGLYHNLSPK 166
 97 PTKQSPLEPML-----DVAEL----- 113
 167 STVDKMLRADVILHKVFLQGYLHLSLNCNGSKSEMIQLGDQEVSELCPREKLA 226
 114 -----LTSLLRT--ESLGALGAQAEPLHSLAEADLAQELLALRSLVELR--A 159
 227 RVLRSNMDILKPIRLTINSTSPFPKELAEATLHLHSLGLTGAQLFMSMSDMR 286
 160 LLQRPRTSGPLEL--LSEALCSVRGSPSTVGPSLNWYEASDLMELVG---QEPE 212
 287 FLTNVSSSSSTQIYQAVSRIVCGHPEGGLKIKSLNWNEDNYKALFGNGTEEDA 346
 213 PDSLSLPACSELIGALDSHPRLRRLLKPLILGKLLFAPDTPFTRKMLAQVNRTPE 272
 347 YDNSTTPYCNLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATQVMAEVNKT 406
 273 TLIRDVREWEMLGPRIFTFMNDSSNVAMLQRLQWQDEGR-RQPRPGGRD----- 322
 407 AVFHDLEGWWEELSPTKITWTFMENSQEMDLVRLMDGRDNDFHFEWQDLDGTD 466
 323 ---HMEALRFLDPGGSG--YSWQDAHADVGHVLTGRVTECLSLDKLEAAPSEAA 377
 467 LAKHPEDVQS-----SNGSVYTWREAFNEFNQAIRISRFWEVCUNLKLPIATE 521
 378 RALQLLAEHRFWAGVVFVLPEDSDPTHEFTPDLPFGPHVRIKIRMDIDVVYTR 437
 522 KSMELDERKFWAGIVFTGTPGSIELPH-----HVYKIRMDIDNVERTNKIDGY 573
 438 WDCGPAADPTLDYRVWGGFVYVQLDLVERAAVRVLSGANPRAGLYLOMPYCVVD 497
 574 WDFGPRADPFEDRMRYVWGGFAYLQDVVEQAIIRVLTGTEKTKTYVMQMPYCV 633
 498 RVLRSRSLPLFLTLAWIYVSTLTAKVAVREKETRLRDMRAMGLSRVAVLWGLF 557
 634 RVMSRSLPFWTLAWIYVSAVIGIYVEKARLKTMRIMGLDNLNLSWFSISLPL 693
 558 LLSAALLVVLKLGDLILPYSHPGVVFVFLAAFAVATVQSFLLSAFFSRANLAA 617
 694 LVSAGLLVWLKLNLLPYSDPSVVFVFLGVFAVAVTILQCLFISTLSFRANLAA 753
 618 YFSLYLPIYVLCVWRDLRPPAGRVASLLSPVAFGCECSIALLEEGEGAHNVG 677
 754 YFTLYLPIYVLCVWQDMYFTLKFISLLSPVAFGCEYFALFEGGIVQWMDNLE 813
 678 T-ADVFLSLAQVSGLLLLDAAALYGLATWYLEAVCPGQYGIPEPWNFPFRSRY 736

CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds. It
CC further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or
CC prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAAL0005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease. Note: The present sequence is not shown
CC in the specification, but is derived from the native human ABC1 shown on
CC pages 152-157

XX Sequence 2261 AA;

Query Match 51.5%; Score 5611.5; DB 3; Length 2261;
Best Local Similarity 50.0%; Pred. No. 0;
Matches 1118; Conservative 361; Mismatches 596; Indels 161; Gaps 26;
QY 1 PPLEHCHFPKPLPSAGTVPMVQLGICNVNNTCFPLQTPGEPGRSLNFDLSVRL 60
DB 47 PPEHCHFPKPLPSAGTVPMVQLGICNVNNTCFPLQTPGEPGRSLNFDLSVRL 106
QY 61 ADARTVGGASAHRTLAGLKLIALTAARSTAO-----PQ 96
DB 107 SDARLLLYSKQTSMDMRKVLRTLOQIKSSSNLKLQDFLVNDFSGFLYHNSLPK 166
QY 97 PTKQSPLEPPML-----DVAEL----- 113
DB 167 STVDKMLRADVLHKVFLQGYQLHLTSLNGSKSEEMQLQDQVSESLCGLPREKLA 226
QY 114 -----LTSLLRT--ESLGLAQOQEPHLSLLEAAEDLAQELLALSLVELR---A 159
DB 227 RVLRNMDILKPIRLTNSTSPPSKELAEATKTLHSLGLTLAGELFSMRSDMRQVM 286
QY 160 LLQPRGTSGPLEL-----LSEALCSVRGSPSTVSGSLNWEASDLMLVG-----QEP 212
DB 287 FLTNVNSSSSTQIYQASRVICVGHPEGGGLIKISLNNYEDNNKALFGNGTEDEAT 346
QY 213 PDSLSLSPACSELICALDHPHLSRLWRLKPLIILGKLLFAPDTPFTRKLAQVNRTEEL 272
DB 347 YDNTTTFYCDMLKMLSESLRSLIWKALKPLIKGLIYPTDPTATQVWAEVNTFQEL 406
QY 273 TLLRDREVWMLGRIFTFNDSNSVAMQLRLLQMDQEGR-RQPRPGGRD----- 322
DB 407 AVFHDLEGWMLSELPKWTFTFMSQEMDLVRMLDSDRDNHFWBQQLDGLDWTQDIVAF 466

QY 323 ---HMEALRSFLDPGSGG--YSWQDAHADVGHVLTGLRVTECLSLDKLEAAPSEAAALVS 377
DB 467 LAKHPEDVQSG-----SNGSVYTWRFAFNETNQAIRTSRMECNLNLKLEPIATEVWLN 521
QY 378 RALQLAAEHRFWAGVVFGLGPESSDPTHEPTDPLGFGHVRKIRMDIDVVYTRTNKIDRF 437
DB 522 KSMELDERKFWAGIVFTGTPGSIELPH-----HVYKIRMDIDVNERTNKIDGY 573
QY 438 WDPGPAADPLTDLRYVWGVFVYLODLVERAAVRLSGANPRAGLIYLOOMPYPCVVDVFL 497
DB 574 WDPGPRADPEDMRYVWGVFVYLODLVERAAVRLSGANPRAGLIYLOOMPYPCVVDVFL 633
QY 498 RVLSRSPLFLTLAWIYSVTLTKAVUREKETRLDTRAMGLSRAVLWGLFSLCLGPF 557
DB 634 RVMSRSMPLFMTLAWIYSVAVIIGVIEKEARKETMRIMGDONSILWFSWFISSLIPL 693
QY 558 LLSAALLVLVLKGLDILPYSHPGVVFLLFAFAFVATVTSFLLSAFSTRANLAACGLA 617
DB 694 LVSAGLLVLVLKGLDILPYSDPSVVFVLSVFAVVTILQCLISTLFSRANLAACGGII 753
QY 618 YFSLYLPYLVCVWRDLRPLAGGRVAASLLSPVAFGFCESLALLEGEQAGAHNVTRP 677
DB 754 YFTLYLPYLVCVWQDVYVGTFLKIFASLLSPVAFGFCGYFALFEQIGVQVMDNLFESP 813
QY 678 T-ADVFSLAQVSGLLLDAAALYGLATWYLEAVCPGYIPEPWNFPRRYSWCGPRPKS 736
DB 814 VEEDGNLTTSVSMMLFDTLYGVMTWYIEAVPGQIGIPRPWFYPTCKYIWGESEDEK 873
QY 737 PAPCPTPLD--PKVLVEEAPPLGSPGVSRLSEKRFPGSPQALRGLSLDFYQGHITAF 795
DB 874 SHFGSNQKRISCEIMEEETHLKLGVSIQNLVKYRDGMKVAVDGLALNFVEGQITSL 933
QY 796 HNGAGKTTLSLUSGLFPSPGGGSAFILGHVDRSMAAIRHLGVCPQVYNVFLDMLTVD 855
DB 934 HNGAGKTTMSILTGLFPPTSGTAYILGDKDIRSEMSTIRQLNGVPCQPHNVLFDMLT 993
QY 856 VMFYGRKLGLSAAVGVPEODRLAQVGL-VSKQSVOTRHLSSGMQKLSVAIAFVGSQV 914
DB 994 IWFYARLGLSEKRVKAEQEMALDVLGSPSSKLSKTSQSLSGMQKLSVALAFVGSKV 1053
QY 915 VILDEPTAGVDPASRRGIWELLKRYREGRTLILSTHLDLAEALIGDRVAVVAGRLCCCG 974
DB 1054 VILDEPTAGVDPSRRGIWELLKRYRGRTIILSTHMDLADVLGDRILAIISHGLCCVG 1113
QY 975 SPLFLRHLGSGYLLILVKARLPLTTN-----EKADTMEGSDVTREKKNKS 1022
DB 1114 SSLFLKNQLGTGYLLTLVKKDVESLSSCRNSSSTVSYLKKEDSVSSQSDAGLSGSD 1173
QY 1023 QGSRVGTFFQLLALVOHWPVCGARLVEELPHLVLVLPYTGADHGSFATLFRLELDRLAEL 1082
DB 1174 DTUTIDVSAISNLRKHVSERLVEDIGHELTVVLPYEAKEGAFVLEFHEIDRLSDLG 1233
QY 1083 LTGYGISDTSLEBIFLKVVEECAADTDMEDGSCQHLCTGIAGLDVTLRLKMPPEQTALE 1142
DB 1234 ISSYGISETTLEBIFLKVAEESGVDAETSDGLTPARRNRRAFG-DKQSCLRPFTEDDAAD 1292
QY 1143 --NGEPAGSAPETDQSGGPDVAG--RVQGWALTRQLOALLKRLFLARSRGLFAQIV 1198
DB 1293 PNDSDIDPESRETDLLSGMDGKGSYQVKGWKLTAQQQVALLKRLFLARSRGLFAQIV 1352
QY 1199 LPALFVGLALVFSILVPPFGHYPALRLSPMYGAQVSFFSESDAPGDPGRARLLLEALQEA 1258
DB 1353 LPAVFVCIALVFSILVPPFGHYPALRLSPMYGAQVSFFSESDAPGDPGRARLLLEALQEA 1412
QY 1259 G-----LBEPPVQHSRHSRFAPEVPAEVAKVLASGNWTPTSPSPACQCSQPGARR 1308
DB 1413 GFGTRCMEGNPIPDTPCQAGEEETWTAPVPQTTIMDLFQNGNWTWQNSPACQCSQDIKK 1472
QY 1309 LLPDCPAAAGPPPPQAQVSGSEVQVNLTCGNLSDFLVKTYPRLVROGLTKKWNVEVRY 1368
DB 1473 MLFVCPFGAGGLPPPPQKQNTADILQDLTGNISDYLKVTYVQIIIAKSLKKNKIWNVEFY 1532
QY 1369 GGFSLG-GRDPLGSLPGQELGRSVEELWALLSPLPGGALDRVLKNTLTAWAHSLSDAQDSLKI 1427


```

||||| 1533 CGFSLGVSNTQALPPSQEVDNAIKQMKHKLAKDSSADRFLNSLGRFWMTGLDTRNNVKV 1592
Qy 1428 WFNNGWHSWAFVNRASNAIIRAHLPDGPBARHAHSITTNHPLNLTKQOLSEAAALMASS 1487
Db 1593 WFNNGWHAISFLVNNAILRANLQGENSPSHGITAFAHPLNLTKQOLSEVALMTTS 1652
Qy 1488 VDVLSICVVFAMSFVPASFTLVLIJEERVTRAKHLQLMGGLSPITLYWLGNFMDMNCNYLV 1547
Db 1653 VDVLSICVVFAMSFVPASFFVFLIQERVSKAHLQFISGVKPVYIWSNFVMDMNCYVV 1712
Qy 1548 PACIVVLFELAPQOQRAYVAPANLPALLLLLLLLYXWSITPLMYPASFPSPSTAYVLTTC 1607
Db 1713 PATLVIIIFICPQKSYVSSTNLPLVALALLLLYXWSITPLMYPASFPKIPSTAYVLTTS 1772
Qy 1608 INLFIGINGMATFVLELFSQKQEVSRILKQVFLIPPHFCLGRGLJDMVRNOAMADAF 1667
Db 1773 VNLFIGINGVAIFVLELFTDNKNININDILKSVFLIPPHFCLGRGLJDMVRNOAMADAL 1832
Qy 1668 ERLGDRQFQSPLRWVVGKNLLIAMVIQGPLFLLTLLQHRSQLLPQPRVRSPLLLGEED 1727
Db 1833 ERFGENRRFVPSLWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDED 1892
Qy 1728 EDVARERERVQAGTQGVVLURNLTKVYRGORMPAVDRLCLGIPPGBCFGLLGVNGAGK 1787
Db 1893 EDVRRERQRIIDGGGQNDILEIKELTKIYRRKRKPAVDRIKIPPGBCFGLLGVNGAGK 1952
Qy 1788 TSTFRMTGDTLASGEAVLAGHSVARBPSSAHLSMGYCQSDAIFELLTGREHLELLAR 1847
Db 1953 SSTFKMLTGDITVTRGDAPLNKNSILSNIEVHQNMGYCQPDATTELLTGREHVEFFAL 2012
Qy 1848 LRGVPEAQVAQTAGSGLARLGLSWYADRPAGTYSGGNKRKLATALALVGDPAVVFLDEPT 1907
Db 2013 LRGVPEKEVGKGEWAIRKGLGVKYEKYNYSNGNKRKLSTANALLIGPPVVFDEPT 2072
Qy 1908 TGMDSARRFLWNSLLAVVREGRSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLK 1967
Db 2073 TGMDSKARRFLWNCALSVVKEGRSVVLTSHSMECEALCTRMAIMVNGRFRCLGSPQHLK 2132
Qy 1968 GFPAAGHTILTLEVPAAARS--QPAARFVAAEFPGSELREAHGGRLPOLPPGRCALARVF 2025
Db 2133 NRFGDGYTIVRIAGSNPDLKPVQDFFGLAFFPGSVLKEKHNMLQYQL-PSLSLARIF 2191
Qy 2026 GELAVHGAEGHVEDFSVQTMLEEVFLYPSKDQKDE--DTEOKEAGVGVDPAPGLQH 2082
Db 2192 SILSQSKRLHIEDYSVQTTLDQVFNFAKDQSDDDHLKDLSTLKNQTV-VDVAV---- 2246
Qy 2083 PKRVSQFLDDPSTAE 2098
Db 2247 ---LTSFLQDEKVES 2259

```

Search completed: January 3, 2005, 10:53:39
Job time : 189.061 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:28:54 ; Search time 108.665 Seconds
(without alignments)
7104.204 Million cell updates/sec

Title: US-09-995-542-5

Perfect score: 11143

Sequence: 1 MAFWTQMLLLKQKFMFYRR.....QHPKRVSQLDPPSTAETVL 2146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11143	100.0	2146	9 US-09-995-542-5	Sequence 5, Appli
2	11130	99.9	2146	15 US-10-182-006-2	Sequence 2, Appli
3	11130	99.9	2180	15 US-10-332-447-27	Sequence 27, Appli
4	11121.5	98.8	2147	17 US-10-618-281-44	Sequence 44, Appli
5	11107	99.7	2144	9 US-09-858-194-2	Sequence 2, Appli
6	11107	99.7	2144	14 US-10-154-419-2	Sequence 2, Appli
7	10896	97.8	2100	9 US-09-995-542-6	Sequence 6, Appli
8	10639.5	95.5	2059	15 US-10-114-270-176	Sequence 176, App
9	9460.5	84.9	1873	15 US-10-182-006-4	Sequence 4, Appli
10	8613.5	77.3	2167	9 US-09-995-542-2	Sequence 2, Appli
11	8389.5	75.3	2121	9 US-09-995-542-3	Sequence 3, Appli
12	7596	68.2	1550	9 US-09-995-542-8	Sequence 8, Appli
13	5768.5	51.8	2261	14 US-10-313-641-9	Sequence 9, Appli

14	5768.5	51.8	2261	14	US-10-313-641-10	Sequence 10, Appli
15	5768.5	51.8	2261	14	US-10-428-551-9	Sequence 9, Appli
16	5768.5	51.8	2261	14	US-10-428-551-10	Sequence 10, Appli
17	5768.5	51.8	2261	15	US-10-452-510-1	Sequence 1, Appli
18	5768.5	51.8	2261	15	US-10-617-334-1	Sequence 1, Appli
19	5768.5	51.8	2261	16	US-10-745-377-5	Sequence 5, Appli
20	5768.5	51.8	2261	16	US-10-744-465-1	Sequence 1, Appli
21	5768.5	51.8	2261	16	US-10-833-679-1	Sequence 1, Appli
22	5768.5	51.8	2261	17	US-10-872-113-5	Sequence 5, Appli
23	5752.5	51.6	2261	10	US-09-984-827-128	Sequence 128, App
24	5752.5	51.6	2261	10	US-09-984-827-132	Sequence 132, App
25	5752.5	51.6	2261	10	US-09-984-827-134	Sequence 134, App
26	5752.5	51.6	2261	10	US-09-984-827-136	Sequence 136, App
27	5751.5	51.6	2261	9	US-09-995-542-11	Sequence 11, Appli
28	5751.5	51.6	2261	9	US-09-845-456-11	Sequence 11, Appli
29	5751.5	51.6	2261	10	US-09-984-827-2	Sequence 2, Appli
30	5751.5	51.6	2261	10	US-09-984-827-127	Sequence 127, App
31	5751.5	51.6	2261	10	US-09-984-827-129	Sequence 129, App
32	5751.5	51.6	2261	10	US-09-984-827-130	Sequence 130, App
33	5751.5	51.6	2261	10	US-09-984-827-131	Sequence 131, App
34	5751.5	51.6	2261	10	US-09-984-827-135	Sequence 135, App
35	5747.5	51.6	2261	14	US-10-340-097-118	Sequence 118, App
36	5747.5	51.6	2261	14	US-10-336-215-118	Sequence 118, App
37	5747.5	51.6	2261	14	US-10-336-219-118	Sequence 118, App
38	5746.5	51.6	2261	10	US-09-984-827-133	Sequence 133, App
39	5744.5	51.6	2263	15	US-10-276-774-2326	Sequence 2326, Ap
40	5531.5	49.6	2201	9	US-09-995-542-9	Sequence 9, Appli
41	5522.5	49.6	2201	14	US-10-170-385-293	Sequence 293, App
42	5522.5	49.6	2201	14	US-10-331-496A-29	Sequence 29, Appli
43	5522.5	49.6	2201	15	US-10-429-160-4	Sequence 4, Appli
44	5522.5	49.6	2201	16	US-10-465-498-139	Sequence 139, App
45	5470.5	49.1	2233	16	US-10-465-498-140	Sequence 140, App

ALIGNMENTS

RESULT 1
US-09-995-542-5
; Sequence 5, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-5

Query Match	100.0%	Score	11143;	DB	9;	Length	2146;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2146;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAFWTQMLLLKQKFMFYRRQP	VOLLV	ELLV	ELLV	ELLV	ELLV
Db	1	MAFWTQMLLLKQKFMFYRRQP	VOLLV	ELLV	ELLV	ELLV	ELLV
Qy	61	LPSAGTVPWLGGLICNVNNTCF	POLT	PGEE	PGRLSNFNDLSV	RLLDARTV	LGASAH
Db	61	LPSAGTVPWLGGLICNVNNTCF	POLT	PGEE	PGRLSNFNDLSV	RLLDARTV	LGASAH
Qy	121	TLAGLGKLIATRAARSTAO	PQTK	SPLE	PPMLDVAELLT	SLRTES	LGALGO
Db	121	TLAGLGKLIATRAARSTAO	PQTK	SPLE	PPMLDVAELLT	SLRTES	LGALGO

Db 121 TLGLGKLIATLRAARSTAQPOKTSQSPLEPPMLDVAELLTSLRTSLGALGQAQEP 180
Qy 181 HSLLEAAEDLAQELLALRSVELRALIQRPRGTSGPLELSEALCSVRGFSSTVGP 240
Db 181 HSLLEAAEDLAQELLALRSVELRALIQRPRGTSGPLELSEALCSVRGFSSTVGP 240
Qy 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDHPRLSRLWRRLKPLILGK 300
Db 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDHPRLSRLWRRLKPLILG 300
Qy 301 TPFTKRLMAQVNRFTFEELTLRDVREYVWMLGPRIFTFMNDSSNAMLQRLQ 360
Db 301 TPFTKRLMAQVNRFTFEELTLRDVREYVWMLGPRIFTFMNDSSNAMLQRLQ 360
Qy 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLE 420
Db 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLE 420
Qy 421 LVSRALQLLAHRFWAGVFLGPEDSSDPTHEPTDLPDGPCHVRKIRMDIDV 480
Db 421 LVSRALQLLAHRFWAGVFLGPEDSSDPTHEPTDLPDGPCHVRKIRMDIDV 480
Qy 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMP 540
Db 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMP 540
Qy 541 VFLRVLSRSLPLFTLAWIYSVTITVKAVREKETRLDRTRAMGLSRAVLWLG 600
Db 541 VFLRVLSRSLPLFTLAWIYSVTITVKAVREKETRLDRTRAMGLSRAVLWLG 600
Qy 601 GPFLLSALLVLKLGDIILPYSPHGVVFLFLAFAVATVTSQFLLSAFSTRAN 660
Db 601 GPFLLSALLVLKLGDIILPYSPHGVVFLFLAFAVATVTSQFLLSAFSTRAN 660
Qy 661 GLAYFSLYLPVLCVAVNRDLRPAAGRVAAALLSPVAFGFCESLALLEBOG 720
Db 661 GLAYFSLYLPVLCVAVNRDLRPAAGRVAAALLSPVAFGFCESLALLEBOG 720
Qy 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPNPFRRSY 780
Db 721 TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYGIPEPNPFRRSY 780
Qy 781 KSPAPCTPLDPKVLVEAPGLSPGVSVRSLEKRPFGSPQALRGSLDFYQGH 840
Db 781 KSPAPCTPLDPKVLVEAPGLSPGVSVRSLEKRPFGSPQALRGSLDFYQGH 840
Qy 841 GHNGAGKTTTILSGLFPSPGSAFILLGHDRVSSMAAIRPHLGVCQYQVNL 900
Db 841 GHNGAGKTTTILSGLFPSPGSAFILLGHDRVSSMAAIRPHLGVCQYQVNL 900
Qy 901 HVMFYGRKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIA 960
Db 901 HVMFYGRKGLSAAVVGPEQDRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIA 960
Qy 961 VILDEPTAGVDPASRGIMWELLKYREGRTLILSTHLLDEAELLGDRVAVAG 1020
Db 961 VILDEPTAGVDPASRGIMWELLKYREGRTLILSTHLLDEAELLGDRVAVAG 1020
Qy 1021 SPLFLRRHLGSGYVLTILVKARLPLTTNEKADTMEGSDVTROEKKNGSQ 1080
Db 1021 SPLFLRRHLGSGYVLTILVKARLPLTTNEKADTMEGSDVTROEKKNGSQ 1080
Qy 1081 LVQHWVFGARLVEBPHLVLVLPYTGADGSGFATLPRELDTRLAELRLTGYS 1140
Db 1081 LVQHWVFGARLVEBPHLVLVLPYTGADGSGFATLPRELDTRLAELRLTGYS 1140
Qy 1141 EIFLKVVEECAADTDMEDGCGQHLCTGIAGLDVTILKWPQPTALENGEPAG 1200
Db 1141 EIFLKVVEECAADTDMEDGCGQHLCTGIAGLDVTILKWPQPTALENGEPAG 1200
Qy 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRSRGLFAQIVLPALFVGLAL 1260
Db 1201 QGSGPDVAVRQGWALTRQQLQALLKRFLLARRSRGLFAQIVLPALFVGLAL 1260

Qy 1261 PPFCHYPALRLSPMTYGAQVSPFSEDA PGDPGRARLLLEALLQAGLEPEP 1320
Db 1261 PPFCHYPALRLSPMTYGAQVSPFSEDA PGDPGRARLLLEALLQAGLEPEP 1320
Qy 1321 PEYPAEVAKVLASGNWTPSPSPACCSQPGARLLPDCPAAAGGPPPOAVT 1380
Db 1321 PEYPAEVAKVLASGNWTPSPSPACCSQPGARLLPDCPAAAGGPPPOAVT 1380
Qy 1381 NLTGRLNSDFLVKTYPRLVROGLTKKVNVEVYGGFSLGGRDPGLSPGQ 1440
Db 1381 NLTGRLNSDFLVKTYPRLVROGLTKKVNVEVYGGFSLGGRDPGLSPGQ 1440
Qy 1441 ALLSPPLGGALDRLVKNLTAWAHSLLDAQDSLKTFWNKNGHSMVAFVNR 1500
Db 1441 ALLSPPLGGALDRLVKNLTAWAHSLLDAQDSLKTFWNKNGHSMVAFVNR 1500
Qy 1501 PGPARHAHSITTLNHPNLTKQOLSEAAALMASSVDVLVSCVVVFAMS 1560
Db 1501 PGPARHAHSITTLNHPNLTKQOLSEAAALMASSVDVLVSCVVVFAMS 1560
Qy 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNLYLPACIVVLI FLAFOQRA 1620
Db 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNLYLPACIVVLI FLAFOQRA 1620
Qy 1621 LLLLLYGWSITPLMYPASFPFSPSTAYVYVLTICINLFIGINGSMATFV 1680
Db 1621 LLLLLYGWSITPLMYPASFPFSPSTAYVYVLTICINLFIGINGSMATFV 1680
Qy 1681 SRLKQVFLTFPHFCILGRGLIDMVRNOAMADAFERLGDROFQSPLRW 1740
Db 1681 SRLKQVFLTFPHFCILGRGLIDMVRNOAMADAFERLGDROFQSPLRW 1740
Qy 1741 GPLFLFTLLLOHRSQLLPQPRVRSPLPGLLEEDVARERERVVQATQ 1800
Db 1741 GPLFLFTLLLOHRSQLLPQPRVRSPLPGLLEEDVARERERVVQATQ 1800
Qy 1801 VYRGQMPAVDRICLIGIPGCEGFLIGVNGAGKTSTFRMVTGDTLASR 1860
Db 1801 VYRGQMPAVDRICLIGIPGCEGFLIGVNGAGKTSTFRMVTGDTLASR 1860
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGRHELELLARLGRVPEAQVACTAG 1920
Db 1861 EPSAAHLSMGYCPQSDAIFELLTGRHELELLARLGRVPEAQVACTAG 1920
Qy 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDDEPTTGMDPSARRFLWNS 1980
Db 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDDEPTTGMDPSARRFLWNS 1980
Qy 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLIRVPAAR 2040
Db 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLIRVPAAR 2040
Qy 2041 EPPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAEGHVEDFSV 2100
Db 2041 EPPGSELREAHGRLRFLQPPGRCALARVFGELAVHGAEGHVEDFSV 2100
Qy 2101 SKDQGDDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAE 2146
Db 2101 SKDQGDDETEQKEAGVGDPAQGLQHPKRVSQFLDDPSTAE 2146

RESULT 2

US-10-182-006-2
; Sequence 2, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23

1
PRIOR APPLICATION NUMBER: PCT/US01/02191
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/177,889
PRIOR FILING DATE: 2000-01-24
PRIOR APPLICATION NUMBER: 60/215,405
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2146
TYPE: PRT
ORGANISM: Homo sapiens
US-10-182-006-2

	Query Match	99.9%	Score 11130	DB 15	Length 2146	
	Best Local Similarity	99.9%	Pred. No. 0			
	Matches 2144	Conservative 1	Mismatches	1	Indels	0
	Gaps	0				
Qy	1	MAFWTQLMLLLWKNFMYRRQPVQQLVLLWFLFLFFILVAVRHSHPPLEHHECHFPNKP	60			
Db	1	MAFWTQLMLLLWKNFMYRRQPVQQLVLLWFLFLFFILVAVRHSHPPLEHHECHFPNKP	60			
Qy	61	LPSAGTVFWLQGLICNVNNTCPQLTGPEEPGRLSNFNDSLSVRLADARTVLCGASAH	120			
Db	61	LPSAGTVFWLQGLICNVNNTCPQLTGPEEPGRLSNFNDSLSVRLADARTVLCGASAH	120			
Qy	121	TLIAGLGLIATLRAARSTAOPTKQSPLEPMLDVAELLTSLRTESLGLAQOEP	180			
Db	121	TLIAGLGLIATLRAARSTAOPTKQSPLEPMLDVAELLTSLRTESLGLAQOEP	180			
Qy	181	HSLLEAAEDLAQELLALRSVLRALLQRPRTSGPSELSALCSVRGSPSTVGP	240			
Db	181	HSLLEAAEDLAQELLALRSVLRALLQRPRTSGPSELSALCSVRGSPSTVGP	240			
Qy	241	YEASDLMELVQEPESALPDSSLSPACSELIGALDSHPLSRLRLKPLILGKLLFAPD	300			
Db	241	YEASDLMELVQEPESALPDSSLSPACSELIGALDSHPLSRLRLKPLILGKLLFAPD	300			
Qy	301	TPFTTKLMAQVNRTEELTLRDREVWMLGPRIFTFMDSSNVAMLRLLQMODEGR	360			
Db	301	TPFTTKLMAQVNRTEELTLRDREVWMLGPRIFTFMDSSNVAMLRLLQMODEGR	360			
Qy	361	QPRPGGRDHEALRSFLDPGSGGYQWQDAHADVGLVGLTGRVTECLSLDKLEAAPSEAA	420			
Db	361	QPRPGGRDHEALRSFLDPGSGGYQWQDAHADVGLVGLTGRVTECLSLDKLEAAPSEAA	420			
Qy	421	LVSRAQLLAEHRFWAGVVFVLPEDSDPTTEHTPDGLGPHVRIKIRMDIDVVTNKR	480			
Db	421	LVSRAQLLAEHRFWAGVVFVLPEDSDPTTEHTPDGLGPHVRIKIRMDIDVVTNKR	480			
Qy	481	DRFWDGGAADPLTLRYVMGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYDD	540			
Db	481	DRFWDGGAADPLTLRYVMGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYDD	540			
Qy	541	VFLVLSRSLPLFLTLAMIYSVTLTVKAVREKETRLDTRAMGLSRAVLMGLWFLSCL	600			
Db	541	VFLVLSRSLPLFLTLAMIYSVTLTVKAVREKETRLDTRAMGLSRAVLMGLWFLSCL	600			
Qy	601	GPFLLSAALLVLVLKGLDILPVSHGVPVFLFLAFAVATVTSQFLLSAFFSRANLAAACG	660			
Db	601	GPFLLSAALLVLVLKGLDILPVSHGVPVFLFLAFAVATVTSQFLLSAFFSRANLAAACG	660			
Qy	661	GLAYESLYPYVLCVAMDRDLPAGGRVAASLLSPVAFGFCESLALLEEQEGAGHNVG	720			
Db	661	GLAYESLYPYVLCVAMDRDLPAGGRVAASLLSPVAFGFCESLALLEEQEGAGHNVG	720			
Qy	721	TRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGQYGIPEPWNFPFRSYWCGPRPP	780			
Db	721	TRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPGQYGIPEPWNFPFRSYWCGPRPP	780			
Qy	781	KSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRGLSLDFYQGHITAPL	840			
Db	781	KSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRGLSLDFYQGHITAPL	840			

Qy	841	GHNGAGKTTTILSILSGLPFPSPGSAFIIIGHDVRSSMAAIRPHLGYCPOVNVFLFDMLTVD	900
Db	841	GHNGAGKTTTILSILSGLPFPSPGSAFIIIGHDVRSSMAAIRPHLGYCPOVNVFLFDMLTVD	900
Qy	901	HWFYGRKLGLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSCGMQRKLSVAIAFVGGSVQ	960
Db	901	HWFYGRKLGLSAAVVGPEQDRLLQDVLGVSKQSVQTRHLSCGMQRKLSVAIAFVGGSVQ	960
Qy	961	VILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEALLGDRVAVVAGGRGCCG	1020
Db	961	VILDEPTAGVDPASRRGIWEILLKYREGRTILSTHHLDEALLGDRVAVVAGGRGCCG	1020
Qy	1021	SPLFRRRLHSGYLLTLVKARLPLTTNEKADTDMEGSVDTQEKKNQSGSVGTGPQLLA	1080
Db	1021	SPLFRRRLHSGYLLTLVKARLPLTTNEKADTDMEGSVDTQEKKNQSGSVGTGPQLLA	1080
Qy	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLPRELDTRLAELRLTGYISDTSL	1140
Db	1081	LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLPRELDTRLAELRLTGYISDTSL	1140
Qy	1141	EIFLKVVEECAADTDMEDGSCGHLCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD	1200
Db	1141	EIFLKVVEECAADTDMEDGSCGHLCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD	1200
Qy	1201	QSGGPDVGRVQGWALTRQQLQALLKRFLLARRRGLFAQIVLPALFVGLALVPSLIV	1260
Db	1201	QSGGPDVGRVQGWALTRQQLQALLKRFLLARRRGLFAQIVLPALFVGLALVPSLIV	1260
Qy	1261	PPFGHYPALRLSPMYGAQVFFSEDAFGDPGRARLLLEALLQEAEGLEPPVQSHSRFSA	1320
Db	1261	PPFGHYPALRLSPMYGAQVFFSEDAFGDPGRARLLLEALLQEAEGLEPPVQSHSRFSA	1320
Qy	1321	PEVPAEVAKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQ	1380
Db	1321	PEVPAEVAKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSEVVQ	1380
Qy	1381	NLTGRNLSDFLVKTVPRLVROGLTKKWNVREVYGGFSLGGDRDCLPGSQELGRSVEELW	1440
Db	1381	NLTGRNLSDFLVKTVPRLVROGLTKKWNVREVYGGFSLGGDRDCLPGSQELGRSVEELW	1440
Qy	1441	ALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIPNNKWHSMVAFVNRASNAILRAHLP	1500
Db	1441	ALLSPLPGCALDRVLKNLTAWAHSIDAQDSLKIPNNKWHSMVAFVNRASNAILRAHLP	1500
Qy	1501	PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSIQVVFAMSVFPASFTLVLTBE	1560
Db	1501	PGPARHAHSITTLNHPNLTKQELSEALMASSVDVLVSIQVVFAMSVFPASFTLVLTBE	1560
Qy	1561	RVTRAKHLQMLGSLPTLYWLNFLWDMCNVLPACIVVLIFLAFQOQRAYVAPANLPALL	1620
Db	1561	RVTRAKHLQMLGSLPTLYWLNFLWDMCNVLPACIVVLIFLAFQOQRAYVAPANLPALL	1620
Qy	1621	LLLLLYGWSITPLMYPASFPFVSPTAYVLTCLNLFINGSMATFVLELFSQKQLEV	1680
Db	1621	LLLLLYGWSITPLMYPASFPFVSPTAYVLTCLNLFINGSMATFVLELFSQKQLEV	1680
Qy	1681	SRILKQVFLIIPPHCLGRGLDMVRNQAMADAPERLGRQFQSPURWEVGNKLLAMVIQ	1740
Db	1681	SRILKQVFLIIPPHCLGRGLDMVRNQAMADAPERLGRQFQSPURWEVGNKLLAMVIQ	1740
Qy	1741	GPLFLLFTLLQHSQSLPQPRVRSPLPGLGEDEVDARERERVOGATGQDVVLVRLNLT	1800
Db	1741	GPLFLLFTLLQHSQSLPQPRVRSPLPGLGEDEVDARERERVOGATGQDVVLVRLNLT	1800
Qy	1801	VYRGQMPAVDRCLGIPPGCEFCGLLVNGAGKSTSTPRMVTGDTLASRGEAVLAGHSVAR	1860
Db	1801	VYRGQMPAVDRCLGIPPGCEFCGLLVNGAGKSTSTPRMVTGDTLASRGEAVLAGHSVAR	1860
Qy	1861	BPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWAD	1920
Db	1861	BPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWAD	1920

[illegible]

RESULT 4

```

RESUL 4
US-10-618-281-44
; Sequence 44, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases

```


Qy 841 GHNGAGKTTTILSILSGLPPSGGSAPILGHDVRSMAAIRPHLGVCPQYVNLFDMLTVDE 900
Db 841 GHNGAGKTTTILSILSGLPPSGGSAPILGHDVRSMAAIRPHLGVCPQYVNLFDMLTVDE 900
Qy 901 HWFYGRLLKGLSAAVGPEQDRLLQDVLVSQSVQTRHLSSGNQKLSVAIAFVGGSQV 960
Db 901 HWFYGRLLKGLSAAVGPEQDRLLQDVLVSQSVQTRHLSSGNQKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGYYTLTKARLPLTTNEKADTMEGSDVTRQEKKNCSQGSRVCTPQLLA 1080
Db 1021 SPLFLRRHLGSGYYTLTKARLPLTTNEKADTMEGSDVTRQEKKNCSQGSRVCTPQLLA 1080
Qy 1081 LVQHWPFCARLVEELPHELVLVLPYTGADHGSFATLFRLELDTRLAELRLTGYSIDTSLE 1140
Db 1081 LVQHWPFCARLVEELPHELVLVLPYTGADHGSFATLFRLELDTRLAELRLTGYSIDTSLE 1140
Qy 1141 EIFLKVVEECAADTMEDGSCGHLCTGAGLDVTLKMPPOBTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTMEDGSCGHLCTGAGLDVTLKMPPOBTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVGRVQGWALTRQOQALLKXFLIARRSRLGFAQIVLPALFVGLALVFSLIV 1260
Db 1201 QGSGPDVGRVQGWALTRQOQALLKXFLIARRSRLGFAQIVLPALFVGLALVFSLIV 1260
Qy 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGDCRARRLEALLQEAAGLEPPVQVHSHRFS 1320
Db 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGDCRARRLEALLQEAAGLEPPVQVHSHRFS 1320
Qy 1321 PEVPAEAKVLASGNWTPESPSPACQSCQARLLPDCPAAAGPPPPQAVTSGSEVQ 1380
Db 1321 PEVPAEAKVLASGNWTPESPSPACQSCQARLLPDCPAAAGPPPPQAVTSGSEVQ 1380
Qy 1381 NLTCGNLSDFLVKTYPRVROGLTKKWNVEVRYGFSGLGRDPCILSGQELGHSVELW 1440
Db 1381 NLTCGNLSDFLVKTYPRVROGLTKKWNVEVRYGFSGLGRDPCILSGQELGHSVELW 1440
Qy 1441 ALLSPLPGCALDRVLKXNLTAHSLDAQDSLKIFWNNKGWSMVAFVNRASNAITLRAHLP 1500
Db 1441 ALLSPLPGCALDRVLKXNLTAHSLDAQDSLKIFWNNKGWSMVAFVNRASNAITLRAHLP 1500
Qy 1501 PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSIICVVFAMSFVPASFTLVLEE 1560
Db 1501 PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSIICVVFAMSFVPASFTLVLEE 1560
Qy 1561 RVTRAKHLQMGSLPTLYLGNFLWDMCNVLPACIVVLIFLAFQORAYVAPANLPA 1620
Db 1561 RVTRAKHLQMGSLPTLYLGNFLWDMCNVLPACIVVLIFLAFQORAYVAPANLPA 1620
Qy 1621 LLLLGYNSITPLMPASFPFSPSTAYVLTCLNLTGINGSMAFVLELFS - QKLOE 1679
Db 1621 LLLLGYNSITPLMPASFPFSPSTAYVLTCLNLTGINGSMAFVLELFS - QKLOE 1680
Qy 1680 VSRILKQVLIFFPHFCILGRGLIDMVRNQAMADAFERLGRDQFQSPLRWEVVGKLLAMVI 1739
Db 1681 VSRILKQVLIFFPHFCILGRGLIDMVRNQAMADAFERLGRDQFQSPLRWEVVGKLLAMVI 1740
Qy 1740 QGPLFLFTLLLOHRSQLLPOPRVRSIPLGEEDEEDVARERERVQGNTOGDVILVNLIT 1799
Db 1741 QGPLFLFTLLLOHRSQLLPOPRVRSIPLGEEDEEDVARERERVQGNTOGDVILVNLIT 1800
Qy 1800 KVRGQMPADVRLCLGIPGCEGFLGVCNAGAKTSTFRMVTGDTLASRGAVALAGHVA 1859
Db 1801 KVRGQMPADVRLCLGIPGCEGFLGVCNAGAKTSTFRMVTGDTLASRGAVALAGHVA 1860
Qy 1860 REPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYA 1919
Db 1861 REPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYA 1920
Qy 1920 DRPAGTYSGGNKKRLATALALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGSV 1979

Db 1921 DRPAGTYSGGNKKRLATALALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGSV 1980
Qy 1980 LTSHSMEECEALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTTLTLRVPAARSQPAAFVA 2039
Db 1981 LTSHSMEECEALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTTLTLRVPAARSQPAAFVA 2040
Qy 2040 AEPFGSELREAHGRRLRFQPLPGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLY 2099
Db 2041 AEPFGSELREAHGRRLRFQPLPGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLY 2100
Qy 2100 FSKDQKDEDETEKQKAGVGDPAQLQHPKRVSQFLDDPSTAEITVL 2146
Db 2101 FSKDQKDEDETEKQKAGVGDPAQLQHPKRVSQFLDDPSTAEITVL 2147

RESULT 5

US-09-858-194-2
; Sequence 2, Application US/09858194
; Patent No. US20020061590A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, MARIA
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF
; FILE REFERENCE: WNI-153
; CURRENT APPLICATION NUMBER: US/09/858,194
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/204,211
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-194-2

Query Match 99.7%; Score 11107; DB 9; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 MAFTQMLMLLWKNFMYRRRQPVQLLVLLWPLFLFFILVAVRSHSPPLSHHECHFPNKP 60
Db 1 MAFTQMLMLLWKNFMYRRRQPVQLLVLLWPLFLFFILVAVRSHSPPLSHHECHFPNKP 58
Qy 61 LPSAGTVPMLOGLICNVNNTCFPOLTPGEPGRLSNPNDSLVSRLLADARTVLGGASAH 120
Db 59 LPSAGTVPMLOGLICNVNNTCFPOLTPGEPGRLSNPNDSLVSRLLADARTVLGGASAH 118
Qy 121 TLAGLGLIATLRAARSTAOPTKOSPLEPPMLDVAELLTSLRTESLGLAQOEPL 180
Db 119 TLAGLGLIATLRAARSTAOPTKOSPLEPPMLDVAELLTSLRTESLGLAQOEPL 178
Qy 181 HSLEAAEDLAQELLALRSILVELRALLQRPRTSGPLELLSEALCSVRGSSSTVGPGLNW 240
Db 179 HSLEAAEDLAQELLALRSILVELRALLQRPRTSGPLELLSEALCSVRGSSSTVGPGLNW 238
Qy 241 YEASDLMELVGPESALPDSISLSPACSELIGALDHSPLSRLLWRRUKPLILGKLLFAPD 300
Db 239 YEASDLMELVGPESALPDSISLSPACSELIGALDHSPLSRLLWRRUKPLILGKLLFAPD 298
Qy 301 TPTRKLMQAVNRTFEELTLRLDRVREWMELGPRIFTFMNDSSNVAMQLQLQMDGRR 360
Db 299 TPTRKLMQAVNRTFEELTLRLDRVREWMELGPRIFTFMNDSSNVAMQLQLQMDGRR 358
Qy 361 QPRPGGRDHMEALRSFLDPCSGGYSWODAHADVHVLGTLGRVTECLSLDKLEAAPEAA 420
Db 359 QPRPGGRDHMEALRSFLDPCSGGYSWODAHADVHVLGTLGRVTECLSLDKLEAAPEAA 418
Qy 421 LVSRALQLLAHFRFWAGVFLGPESSDPTHEHTPDLPQGHVRIKIRMDIDVVTTRTKIR 480
Db 419 LVSRALQLLAHFRFWAGVFLGPESSDPTHEHTPDLPQGHVRIKIRMDIDVVTTRTKIR 478

QY 481 DRFWDPGPAADPLTDLRYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYVDD 540
DB 479 DRFWDPGPAADPLTDLRYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQMPYCYVDD 538
QY 541 VFLRVLSRLPLFLTLAWIYSVTLTAKAVVREKETRLRDTMRAMGLSRAVLWLGWFLSCL 600
DB 539 VFLRVLSRLPLFLTLAWIYSVTLTAKAVVREKETRLRDTMRAMGLSRAVLWLGWFLSCL 598
QY 601 GPFLLSAALLVLVLKGLDILPYSHPGVWFLFLAAFAVATVTQSFLLSAFFSRAANLAAACG 660
DB 599 GPFLLSAALLVLVLKGLDILPYSHPGVWFLFLAAFAVATVTQSFLLSAFFSRAANLAAACG 658
QY 661 GLAYFSLYLPYVLCVAMWRDLRPAAGRVAASLLSPVAFGFCESLALLBEEQGEAOWHNVG 720
DB 659 GLAYFSLYLPYVLCVAMWRDLRPAAGRVAASLLSPVAFGFCESLALLBEEQGEAOWHNVG 718
QY 721 TRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGQYIPEPWNFFPRRSYWCGRPP 780
DB 719 TRPTADVFSLAQVSGLLLLDAAALYGLATWYLEAVCPGQYIPEPWNFFPRRSYWCGRPP 778
QY 781 KSPAPCPTPLDPKVLVEAPGLSPQSVRSLEKFPGPSQALRGLSLDFYQGHITAPL 840
DB 779 KSPAPCPTPLDPKVLVEAPGLSPQSVRSLEKFPGPSQALRGLSLDFYQGHITAPL 838
QY 841 GHNGAGKTTTSLSGLPPSPGGSFAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTUDE 900
DB 839 GHNGAGKTTTSLSGLPPSPGGSFAFILGHVDRSSMAAIRPHLGVCPQYNVLFDMLTUDE 898
QY 901 HWYFVGRUKGLSAAVVGPEQDRLLQDVLVSKQSVOTRHLSCGMQRKLSVAIFVGGSQV 960
DB 899 HWYFVGRUKGLSAAVVGPEQDRLLQDVLVSKQSVOTRHLSCGMQRKLSVAIFVGGSQV 958
QY 961 VILDEPTAGVDPASRRGTWELLKVKREGRTLLSTHILDEALLGDRVAVVAGRLCCCG 1020
DB 959 VILDEPTAGVDPASRRGTWELLKVKREGRTLLSTHILDEALLGDRVAVVAGRLCCCG 1018
QY 1021 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKNKSGQSRVGTVPQLLA 1080
DB 1019 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKNKSGQSRVGTVPQLLA 1078
QY 1081 LVQHWVPGARLVEELPHELVLVLVLPYTGADHGSFATLPRELDRLAELRLTGVISDTSLE 1140
DB 1079 LVQHWVPGARLVEELPHELVLVLVLPYTGADHGSFATLPRELDRLAELRLTGVISDTSLE 1138
QY 1141 EFLKVVBECAADTMEGSCGQHLCGTAGLDVTLRLKMPQOETALENGEPAGSAPETD 1200
DB 1139 EFLKVVBECAADTMEGSCGQHLCGTAGLDVTLRLKMPQOETALENGEPAGSAPETD 1198
QY 1201 QGSGPDVAVRGVQWALTRQQLALLKRFLLARRRRGLFAQIVLPALFVGLALVFSLIV 1260
DB 1199 QGSGPDVAVRGVQWALTRQQLALLKRFLLARRRRGLFAQIVLPALFVGLALVFSLIV 1258
QY 1261 PPGHYPALRLSPTMYGAQVSPFSDADPGDGRARLLEALLQEALEBPVQVHSHRPSA 1320
DB 1259 PPGHYPALRLSPTMYGAQVSPFSDADPGDGRARLLEALLQEALEBPVQVHSHRPSA 1318
QY 1321 PVPVPAEVAKVLASGNWTPESPACQCPGARRLLPDCPAAAGGPPPPQAVTGSGEVVQ 1380
DB 1319 PVPVPAEVAKVLASGNWTPESPACQCPGARRLLPDCPAAAGGPPPPQAVTGSGEVVQ 1378
QY 1381 NLTGRLNSDFLVKTYTFLRVROGLKTKKWVNEVYGGFSLGGDRDGLPSQELGRSVERLW 1440
DB 1379 NLTGRLNSDFLVKTYTFLRVROGLKTKKWVNEVYGGFSLGGDRDGLPSQELGRSVERLW 1438
QY 1441 ALLSPLPGALDRVLKNTTAWAHSIDAQDSLKIWFNNKGWHSWAFVNRASNAIIRAHLP 1500
DB 1439 ALLSPLPGALDRVLKNTTAWAHSIDAQDSLKIWFNNKGWHSWAFVNRASNAIIRAHLP 1498
QY 1501 PGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVVPFAMSFPASFTLVLEE 1560
DB 1499 PGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVVPFAMSFPASFTLVLEE 1558
QY 1561 RVTRAKHLQLMGGLSPTLYLWGNFLWDMCNVLPACIVVLIFLAFQQRAYVAPANLPALL 1620

DB 1559 RVTRAKHLQLMGGLSPTLYLWGNFLWDMCNVLPACIVVLIFLAFQQRAYVAPANLPALL 1618
QY 1621 LLLLYLXGHSITPLMYPASFFFSVPSTAYVVLTCINLFTGINGSMATFVLELFSQDKLQBV 1680
DB 1619 LLLLYLXGHSITPLMYPASFFFSVPSTAYVVLTCINLFTGINGSMATFVLELFSQDKLQBV 1678
QY 1681 SRLKQVFLIIPPHFCLGRGLIDMVNRQAMADAFERLGRQFQFOSPLURWEVVGKLLAMVQ 1740
DB 1679 SRLKQVFLIIPPHFCLGRGLIDMVNRQAMADAFERLGRQFQFOSPLURWEVVGKLLAMVQ 1738
QY 1741 GPLFLTLTLQLHRSQLLPQPRVRSPLPLGEEDEDVARERERVVQATQGDVVLRLNLTK 1800
DB 1739 GPLFLTLTLQLHRSQLLPQPRVRSPLPLGEEDEDVARERERVVQATQGDVVLRLNLTK 1798
QY 1801 VYRGQMPAVDRCLCLGIPPGECFGLLGVNGAGKTSITFRMVTGDTLASGEAVLAGHSVAR 1860
DB 1799 VYRGQMPAVDRCLCLGIPPGECFGLLGVNGAGKTSITFRMVTGDTLASGEAVLAGHSVAR 1858
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVAGTAGSLARLGLSWAD 1920
DB 1859 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVAGTAGSLARLGLSWAD 1918
QY 1921 RPAGTYSGNKRKLATALALVGDPAVVFDEPTTGMDFPSARRFLWNSLLAVVREGSVML 1980
DB 1919 RPAGTYSGNKRKLATALALVGDPAVVFDEPTTGMDFPSARRFLWNSLLAVVREGSVML 1978
QY 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAFVAA 2040
DB 1979 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAARSQPAAFVAA 2038
QY 2041 EFPQSELREAHGGRRLRFOLPPGRCALARVFGELAVHCAEHGVEDFSVQTMLEBFLYF 2100
DB 2039 EFPQSELREAHGGRRLRFOLPPGRCALARVFGELAVHCAEHGVEDFSVQTMLEBFLYF 2098
QY 2101 SKDQCKDDETEQKEAGVGDVPAPGLQHPKRVSQFLDDPSTAETVL 2146
DB 2099 SKDQCKDDETEQKEAGVGDVPAPGLQHPKRVSQFLDDPSTAETVL 2144

RESULT 6

US-10-154-419-2
; Sequence 2, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:

; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58293FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154.419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 2
; LENGTH: 2144
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-154-419-2

Query Match 99.7%; Score 11107; DB 14; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 MAFVTQMLLLKWKPMYRRROPVOLLVLELLWPLFLFFTLVAVRSHHPLEHHECHFPNKP 60
DB 1 MAFVTQMLLLKWKPMYRRROPVOLLVLELLWPLFLFFTLVAVRSHHPLEHHECHFPNKP 58
QY 61 LPSAGTVPLWQGLICNVNNTCFPQITPGEEPGRLSNFNDSLVSRLADARTVLGGASAH 120

Db 59 LPSAGTVPWLOGLICNVNNTCFPQTPGEEPCRLSNFNDSLVSRLLDARTVLGCASAH 118
Qy 121 TLAGLGLKIATRAARSTAOPTKQSPLEPPMLDVAELTSLTSLRTESLGLAQOEPL 180
Db 119 TLAGLGLKIATRAARSTAOPTKQSPLEPPMLDVAELTSLTSLRTESLGLAQOEPL 178
Qy 181 HSLEAAEDLAQELLALRSVLVELRALLORPGTSGPLELSEALCSVRGSPSTVGPISLNW 240
Db 179 HSLEAAEDLAQELLALRSVLVELRALLORPGTSGPLELSEALCSVRGSPSTVGPISLNW 238
Qy 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDHPLSRLWRRLLKPLILGKLLFAPD 300
Db 239 YEASDLMELVQEPESALPDSLSLSPACSELIGALDHPLSRLWRRLLKPLILGKLLFAPD 298
Qy 301 TPFTRKMAQVNRTFEELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLLOQDEGRR 360
Db 299 TPFTRKMAQVNRTFEELTLRDVREVMELGPRIFTFMNDSSNVAMLQRLLOQDEGRR 358
Qy 361 QPRGGRDHMBALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLEAAPSEAA 420
Db 359 QPRGGRDHMBALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLEAAPSEAA 418
Qy 421 LVSRALQLLAHRFWAGVFLGPEDSDPTHEPTPDIGPGHVRIRKIMDIDVTRTKIR 480
Db 419 LVSRALQLLAHRFWAGVFLGPEDSDPTHEPTPDIGPGHVRIRKIMDIDVTRTKIR 478
Qy 481 DRFPDGPADPLTDLRYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPYCVD 540
Db 479 DRFPDGPADPLTDLRYVMGFFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPYCVD 538
Qy 541 VFLRVLSRLPLFLTLAWIYSVTITVKAUVREKTRLDTRWANGLSRAVLWLGWFLSCL 600
Db 539 VFLRVLSRLPLFLTLAWIYSVTITVKAUVREKTRLDTRWANGLSRAVLWLGWFLSCL 598
Qy 601 GPFLLSAALLVLKIGDILPYSHPGVVFLLAFAFATVTTQSLLSAFTFSRANLAAACG 660
Db 599 GPFLLSAALLVLKIGDILPYSHPGVVFLLAFAFATVTTQSLLSAFTFSRANLAAACG 658
Qy 661 GLAYFSLYLPLYVLCVAMRDLRPPAGGRVAASLLSPVAFGFCESLALLEQEGEAQWNVG 720
Db 659 GLAYFSLYLPLYVLCVAMRDLRPPAGGRVAASLLSPVAFGFCESLALLEQEGEAQWNVG 718
Qy 721 TRPTADVLSLAQSGLLLLDAAVLGLATWLEAVCPQGYGIPBPWNPFFRRSYWCGRRPP 780
Db 719 TRPTADVLSLAQSGLLLLDAAVLGLATWLEAVCPQGYGIPBPWNPFFRRSYWCGRRPP 778
Qy 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRGHSLDFYQGHITAF 840
Db 779 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRFPGSPQALRGHSLDFYQGHITAF 838
Qy 841 GHNGAGKTTTLLSILSGLFPSPGSAFILGHHDVRSMAAIRPHLGVCQYXNVLFDMLTVD 900
Db 839 GHNGAGKTTTLLSILSGLFPSPGSAFILGHHDVRSMAAIRPHLGVCQYXNVLFDMLTVD 898
Qy 901 HWYFYGRILKGLSAAVGPEQDRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIAFVGGSQV 960
Db 899 HWYFYGRILKGLSAAVGPEQDRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIAFVGGSQV 958
Qy 961 VILDEPTAGVDPASRRGIWELLKRYRGRTLILSTHLLDBAELLGDRVAVVAGRLCCCG 1020
Db 959 VILDEPTAGVDPASRRGIWELLKRYRGRTLILSTHLLDBAELLGDRVAVVAGRLCCCG 1018
Qy 1021 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTQEKKNGSQGSRVGTPOLLA 1080
Db 1019 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTQEKKNGSQGSRVGTPOLLA 1078
Qy 1081 LVQHWVPGARLVEELPHELVLVILPYTCANDGSPATLFLRELDTRLAELRLTGYGIDSPTSLE 1140
Db 1079 LVQHWVPGARLVEELPHELVLVILPYTCANDGSPATLFLRELDTRLAELRLTGYGIDSPTSLE 1138
Qy 1141 EIFLKVVBECAADTDMEDGSCQHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETD 1200

Db 1139 EIFLKVVBECAADTDMEDGSCQHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETD 1198
Qy 1201 QSGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGPAQIVLPALFVGLALVFSLI 1260
Db 1199 QSGSPDAVGRVQGWALTRQOLQALLKRFLLARRSRRLGPAQIVLPALFVGLALVFSLI 1258
Qy 1261 PPGHYPALRLSPMTYGAQVSFFSEADPGDPGRARLLEALLOEAGLEEPVQVSHSRFSA 1320
Db 1259 PPGHYPALRLSPMTYGAQVSFFSEADPGDPGRARLLEALLOEAGLEEPVQVSHSRFSA 1318
Qy 1321 PEYPAEVAKVLASGNWTPESPACQCSQPGARLLPDCPAAAGGPPPPQAVTGSGEVQ 1380
Db 1319 PEYPAEVAKVLASGNWTPESPACQCSQPGARLLPDCPAAAGGPPPPQAVTGSGEVQ 1378
Qy 1381 NLGRLNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELW 1440
Db 1379 NLGRLNLSDFLVKTYPRLVQGLTKKWNVEVYGGFSLGGRDPGLPSGQELGRSVEELW 1438
Qy 1441 ALLSPLPGGALDRVLKNTAWAHSLLDAODSLKTFWNNKGWHSVAFVNRASNAILLRAHLP 1500
Db 1439 ALLSPLPGGALDRVLKNTAWAHSLLDAODSLKTFWNNKGWHSVAFVNRASNAILLRAHLP 1498
Qy 1501 PGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSI CVVPFAMSFVPASFTLVLEE 1560
Db 1499 PGPARHAHSITTLNHPNLNTKEQLSEAAALMASSVDVLVSI CVVPFAMSFVPASFTLVLEE 1558
Qy 1561 RVTRAKHLQMGGLSPTLYWLGFLWDMCNLYVPACIVLILFLAPOORAVVAPANLPALL 1620
Db 1559 RVTRAKHLQMGGLSPTLYWLGFLWDMCNLYVPACIVLILFLAPOORAVVAPANLPALL 1618
Qy 1621 LLLLLVGWSTITPLMYPASPEFSVPSTAYVVLTCINLFIGINGSMATFVLFSQDKLOEV 1680
Db 1619 LLLLLVGWSTITPLMYPASPEFSVPSTAYVVLTCINLFIGINGSMATFVLFSQDKLOEV 1678
Qy 1681 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVGKLLAMVIQ 1740
Db 1679 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPRLRWEVVGKLLAMVIQ 1738
Qy 1741 GPLFLFTLLQLHRSQLLPQPRVRSPLPGEEDVARERERVVQATQGDVILVRLNLT 1800
Db 1739 GPLFLFTLLQLHRSQLLPQPRVRSPLPGEEDVARERERVVQATQGDVILVRLNLT 1798
Qy 1801 VYRGORMPAVDRCLGIPGCECFGLGVNGAGKTSTFRMVTGDTLASRGEAVLAGHSVAR 1860
Db 1799 VYRGORMPAVDRCLGIPGCECFGLGVNGAGKTSTFRMVTGDTLASRGEAVLAGHSVAR 1858
Qy 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGSWYAD 1920
Db 1859 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGSWYAD 1918
Qy 1921 RPAGTYSGGNKRKLATALAVGDPVAVVFLDDEPTTGMDDPSARRFLMNSLLAVVREGRSVML 1980
Db 1919 RPAGTYSGGNKRKLATALAVGDPVAVVFLDDEPTTGMDDPSARRFLMNSLLAVVREGRSVML 1978
Qy 1981 TSHSMECEALCSRLAIVNNGRFRCLGSPHOLKGRFAAGHTLTLRVPAASQAAAAFVAA 2040
Db 1979 TSHSMECEALCSRLAIVNNGRFRCLGSPHOLKGRFAAGHTLTLRVPAASQAAAAFVAA 2038
Qy 2041 EFPGESELRHAHGRRLRFLQPLPGGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYF 2100
Db 2039 EFPGESELRHAHGRRLRFLQPLPGGRCALARVFGELAVHGAHGVDEFSVQTMLEEVFLYF 2098
Qy 2101 SKDQKDEDETEQKEAGVGVDPAQLOHPKRVSQFLLDDPSTAEVT 2146
Db 2099 SKDQKDEDETEQKEAGVGVDPAQLOHPKRVSQFLLDDPSTAEVT 2144

RESULT 7
US-09-995-542-6
; Sequence 6, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John

APPLICANT: Ulias, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT PILING DATE: 2001-11-28
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-6

Query Match 97.8%; Score 10896; DB 9; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	47	PLEHHECHPPNKPPLPSAGTVPWLOGLICNVNNTCEPOLTRGEEPCRLSNFNDLSVRL	106
DB	1	PLEHHECHPPNKPPLPSAGTVPWLOGLICNVNNTCEPOLTRGEEPCRLSNFNDLSVRL	60
QY	107	ADARTVLGGASAHRTLAGLGLKIATLRAARSTAQPOPTKQSPLEPPMLDVAELLTSL	166
DB	61	ADARTVLGGASAHRTLAGLGLKIATLRAARSTAQPOPTKQSPLEPPMLDVAELLTSL	120
QY	167	ESLGALGQAQBPPLSHLLBAEDLAQELLALSLVELRALLQRGTSGPLELSEALCS	226
DB	121	ESLGALGQAQBPPLSHLLBAEDLAQELLALSLVELRALLQRGTSGPLELSEALCS	180
QY	227	VRGPSTVGPISNWEASDLMELVGOEPESALPDSSLSPACSELICALDHPSLRWR	286
DB	181	VRGPSTVGPISNWEASDLMELVGOEPESALPDSSLSPACSELICALDHPSLRWR	240
QY	287	LKPLILGKLLFADPTPFTKRLMAQVNRITPEELTLRDREVWMLGPRIFTFMDSSNVA	346
DB	241	LKPLILGKLLFADPTPFTKRLMAQVNRITPEELTLRDREVWMLGPRIFTFMDSSNVA	300
QY	347	MLQRLQWQDEGRORPGRDHMEALRSFLDPGSGYSWQADHVLVGLTGRVTEC	406
DB	301	MLQRLQWQDEGRORPGRDHMEALRSFLDPGSGYSWQADHVLVGLTGRVTEC	360
QY	407	LSLDKLEAPSEALVSALQALLAHRFWAGVVLFGPEDSSDPTHEPTDPLGPGHVRKI	466
DB	361	LSLDKLEAPSEALVSALQALLAHRFWAGVVLFGPEDSSDPTHEPTDPLGPGHVRKI	420
QY	467	RMDIDVTRTNKIRDRFWDPPGAADPLTDLRYVMGGFVYLDLVERAAVRVLSGANPRAG	526
DB	421	RMDIDVTRTNKIRDRFWDPPGAADPLTDLRYVMGGFVYLDLVERAAVRVLSGANPRAG	480
QY	527	LYLOQMPYCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDRTWAMGL	586
DB	481	LYLOQMPYCYVDDVFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKETRLDRTWAMGL	540
QY	587	SRVNLWGLFSLCIGPFLSALVILVILKGLDILPYSHPGVVLFLAFAFVATVTSFLL	646
DB	541	SRVNLWGLFSLCIGPFLSALVILVILKGLDILPYSHPGVVLFLAFAFVATVTSFLL	600
QY	647	SAFFSRANLAAACGGGLAFSLYPLVVLVAVWRDLRPAAGRVAASLLSPVAFGFCESLAL	706
DB	601	SAFFSRANLAAACGGGLAFSLYPLVVLVAVWRDLRPAAGRVAASLLSPVAFGFCESLAL	660
QY	707	LEEQEGEAQWNVGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPGQYGIPEPN	766
DB	661	LEEQEGEAQWNVGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPGQYGIPEPN	720
QY	767	PPFRSYWCGPRPPKSPACPTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPQPALRG	826
DB	721	PPFRSYWCGPRPPKSPACPTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPQPALRG	780
QY	827	LSLDFYQGHITAFLGHNGAGKTTTLLSLSGLFPPSGGSFAFILGHVDRSSMAAIRPHLVC	886

DB	781	LSLDFYQGHITAFLGHNGAGKTTTLLSLSGLFPPSGGSFAFILGHVDRSSMAAIRPHLVC	840
QY	887	POYNVLFDMLTVDHEWVFYGRUKGLSAAVGPEDQRLLDQVLGVSKQSVQTRHLSGMQR	946
DB	841	POYNVLFDMLTVDHEWVFYGRUKGLSAAVGPEDQRLLDQVLGVSKQSVQTRHLSGMQR	900
QY	947	KLSVAIAFVGSSQVILDEPTAGVDPASRRGTWELLKYREGRTILSTHHLDEARELLGD	1006
DB	901	KLSVAIAFVGSSQVILDEPTAGVDPASRRGTWELLKYREGRTILSTHHLDEARELLGD	960
QY	1007	RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN	1066
DB	961	RVAVVAGRLCCCGSPLFLRRHLGSGYYLTLVKARLPLTTNEKADTDMEGSVDTROEKN	1020
QY	1067	SGQSRVGTPLLALVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLFBELDTRLA	1126
DB	1021	SGQSRVGTPLLALVQHWVPGARLVEBELPHELVLVLPYTGADHGSFATLFBELDTRLA	1080
QY	1127	LRLTCYGISDTSLEEIFLKVVVEECAADTDMEDGSCQHLCCTGCIAGLDVTLRLKMPQETA	1186
DB	1081	LRLTCYGISDTSLEEIFLKVVVEECAADTDMEDGSCQHLCCTGCIAGLDVTLRLKMPQETA	1140
QY	1187	LENGEPAASAPETDOGSGPDVGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP	1246
DB	1141	LENGEPAASAPETDOGSGPDVGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP	1200
QY	1247	ALFVGLALVLSLIVPPFGHYPALRLSPMYGAQVSPFSEDPGDPGRARLLLEALLQEA	1306
DB	1201	ALFVGLALVLSLIVPPFGHYPALRLSPMYGAQVSPFSEDPGDPGRARLLLEALLQEA	1260
QY	1307	BEPPVOHSHRFSAPAEVPAEVAKVLASGNWTPESPSPACQSQPCARLLPDCPAAAGP	1366
DB	1261	BEPPVOHSHRFSAPAEVPAEVAKVLASGNWTPESPSPACQSQPCARLLPDCPAAAGP	1320
QY	1367	PPQAVTSGGEVQNLGTGRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGPSLGRDPL	1426
DB	1321	PPQAVTSGGEVQNLGTGRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGPSLGRDPL	1380
QY	1427	PSGQELGRSVBELWALLSPLPGGALDRVLKNTAWAHSLSDAQSLKIFWNNKGMHSMVAF	1486
DB	1381	PSGQELGRSVBELWALLSPLPGGALDRVLKNTAWAHSLSDAQSLKIFWNNKGMHSMVAF	1440
QY	1487	VNRASNALRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVVFAM	1546
DB	1441	VNRASNALRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSIQVVFAM	1500
QY	1547	SFVPASFVLVTEBRTVRAKHLQLMGGLSPTLYWLNFLWDMCNVLPACIVVLIPLAFQ	1606
DB	1501	SFVPASFVLVTEBRTVRAKHLQLMGGLSPTLYWLNFLWDMCNVLPACIVVLIPLAFQ	1560
QY	1607	QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT	1666
DB	1561	QRAYVAPANLALLLLLYGWSITPLMYPASFFSVSTAYVVLTCINLFIGINGSMAT	1620
QY	1667	FVLELFSQKLOEVSRIILKQVFLIFPHFCLGRGLIDWVRNOAMADAFERLGRQOSPLR	1726
DB	1621	FVLELFSQKLOEVSRIILKQVFLIFPHFCLGRGLIDWVRNOAMADAFERLGRQOSPLR	1680
QY	1727	WEVVGKLLAMVIOGPIFLFTLLLOHRSQLLPOPRVRSPLPGLGEEDVARERERVQ	1786
DB	1681	WEVVGKLLAMVIOGPIFLFTLLLOHRSQLLPOPRVRSPLPGLGEEDVARERERVQ	1740
QY	1787	ATQGDVLVRLNLTKVYRGQRMPAVDRCLGIPPGCEGGLGVNGAGKTTSTFRMTVGTDLA	1846
DB	1741	ATQGDVLVRLNLTKVYRGQRMPAVDRCLGIPPGCEGGLGVNGAGKTTSTFRMTVGTDLA	1800
QY	1847	SRGAVALAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRGPEAQVATA	1906
DB	1801	SRGAVALAGHSVAREPSAAHLSMGYCPOSDAIFELLTGREHLELLARLGRGPEAQVATA	1860
QY	1907	GSGLARLGLSWADRPAGTYSGGNKRKATALAVGDPVAVFLDEPTTGMPSARRFLWN	1966

Db 1861 GSGIARLGSLWYADRPAGTSGGNKRKLTALALVGDPAVVFLDEPTTGMDPSARRFLWN 1920
Qy 1967 SLLAVVREGSVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 2026
Db 1921 SLLAVVREGSVMLTSHSMEECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTTLRV 1980
Qy 2027 PAARSQAPAAFAVAEFPGSELREAHGGRRLRFPOLPPGRCALARVFGELAVHGAHGVDEF 2086
Db 1981 PAARSQAPAAFAVAEFPGSELREAHGGRRLRFPOLPPGRCALARVFGELAVHGAHGVDEF 2040
Qy 2087 SVSOTMLEEFLYPSKQGDDETEEOKEAGVGVDPAAGLQHPKRVSQFLDDPSTAETVL 2146
Db 2041 SVSOTMLEEFLYPSKQGDDETEEOKEAGVGVDPAAGLQHPKRVSQFLDDPSTAETVL 2100

RESULT 8

US-10-114-270-176
; Sequence 176, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaozhong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenna
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234

; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 176
; LENGTH: 2059
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-176

Query Match 95.5%; Score 10639.5; DB 15; Length 2059;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2058; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 MAFWTQLMLLLWKNFMYRRRQPVQLLVLELLWPLFLFPILVAVRSHHPPLHHECHFPNKP 60
Db 1 MAFWTQLMLLLWKNFMYRRRQPVQLLVLELLWPLFLFPILVAVRSHHPPLHHECHFPNKP 60
Qy 61 LPSAGTVPMLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPMLQGLICNVNNTCFPOLTPGEBPGRLSNFNDSLVSRLLADARTVLGGASAH 120
Qy 121 TLAGLGKLIATLRAARSTAOPTKOSPLEPPMLDVAELTSLRTESLGALGQAEPL 180
Db 121 TLAGLGKLIATLRAARSTAOPTKOSPLEPPMLDVAELTSLRTESLGALGQAEPL 180
Qy 181 HSLLAAEDIAQELLALRSIVELRALLQRPRTSGPLELLSEALCSVRGSSSTVGPSSLNW 240
Db 181 HSLLAAEDIAQELLALRSIVELRALLQRPRTSGPLELLSEALCSVRGSSSTVGPSSLNW 240
Qy 241 YEASDLMELVQEPESALPDSSISPACSELIIGALDHPLSRLLRRLKPLILGKLLFAPD 300
Db 241 YEASDLMELVQEPESALPDSSISPACSELIIGALDHPLSRLLRRLKPLILGKLLFAPD 300
Qy 301 TPFTKRLMAQVNRTEELTLRLDRVREYWEMLGPRIFTFMNDSNVAMLRLLQMDDEGRR 360
Db 301 TPFTKRLMAQVNRTEELTLRLDRVREYWEMLGPRIFTFMNDSNVAMLRLLQMDDEGRR 360
Qy 361 QRPGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Db 361 QRPGRDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Qy 421 LVSRALQLLAHRFWAGVFLGPEDSDPTHEHTPDLPGRHVRKIRMDIDVVRTNKIR 480
Db 421 LVSRALQLLAHRFWAGVFLGPEDSDPTHEHTPDLPGRHVRKIRMDIDVVRTNKIR 480
Qy 481 DRFWDPGFAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
Db 481 DRFWDPGFAADPLTLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQMPYPCYVDD 540
Qy 541 VFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKTRLDTRMRANGLSRAVLWGLWFLSCL 600
Db 541 VFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKTRLDTRMRANGLSRAVLWGLWFLSCL 600
Qy 601 GPFLLSAALVLVLKGLDILPYSHPGVFLFLAFAVATVTSFLLSAFTSRANLAAACG 660
Db 601 GPFLLSAALVLVLKGLDILPYSHPGVFLFLAFAVATVTSFLLSAFTSRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVWRDLRPAAGGRVAASLLSPVAFPGCESIALLEEGEGAQWNVG 720
Db 661 GLAYFSLYLPVLCVWRDLRPAAGGRVAASLLSPVAFPGCESIALLEEGEGAQWNVG 720
Qy 721 TRPTADVFLSAQVSGLLLLDAALYGLATWTLVAVCPQYGIPEPWNPFRRSYWCGRRPP 780
Db 721 TRPTADVFLSAQVSGLLLLDAALYGLATWTLVAVCPQYGIPEPWNPFRRSYWCGRRPP 780
Qy 781 KSPAPCPTLPDVKVLEVEAPPGLSPGVSRSLKRFPGSPQALRGLSLDFYQGHITAF 840
Db 781 KSPAPCPTLPDVKVLEVEAPPGLSPGVSRSLKRFPGSPQALRGLSLDFYQGHITAF 840
Qy 841 GHNGAGKTTTSLISGLFPFGSGSAFLLGHHDVRSMAAIIIRPHLVCQYQNVLFMLTVDE 900
Db 841 GHNGAGKTTTSLISGLFPFGSGSAFLLGHHDVRSMAAIIIRPHLVCQYQNVLFMLTVDE 900

QY 901 HWFYGRKGLSAAVVGPEQDRLLQDVLVSKSQSVQTRHLSGGMQRKLSVAIAFVGGSQV 960
Db 901 HWFYGRKGLSAAVVGPEQDRLLQDVLVSKSQSVQTRHLSGGMQRKLSVAIAFVGGSQV 960
QY 961 VILDEPTAGVDPASRRGTWELLKVKRGRITLSTHLLDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGTWELLKVKRGRITLSTHLLDEALLGDRVAVVAGRLCCCG 1020
QY 1021 SPLFURRLHSGSYILTLVKARLPLTTNEKADTDMEGSVDTROEKKNGSQSGSVGTQPQLLA 1080
Db 1021 SPLFURRLHSGSYILTLVKARLPLTTNEKADTDMEGSVDTROEKKNGSQSGSVGTQPQLLA 1080
QY 1081 LVQHWVPGARLVEELPHELVLPVYTGADHGSFATLFRDLTRIAELRLTGYISDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLPVYTGADHGSFATLFRDLTRIAELRLTGYISDTSLE 1140
QY 1141 EFLKVVCECAADTMEGSCCOHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
Db 1141 EFLKVVCECAADTMEGSCCOHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
QY 1201 QSGSDPAVGRVOGWALTRQQLQALLKRFLIARRRRGLFAOIVLPALFVGLALVPSLIV 1260
Db 1201 QSGSDPAVGRVOGWALTRQQLQALLKRFLIARRRRGLFAOIVLPALFVGLALVPSLIV 1260
QY 1261 PPGHYPALRLSPWTMYGAQVSFSDAPGDPGRARLEALLQEAAGLEPPVQHSRRFSA 1320
Db 1261 PPGHYPALRLSPWTMYGAQVSFSDAPGDPGRARLEALLQEAAGLEPPVQHSRRFSA 1320
QY 1321 PEVPAEVAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGPPPPQAVTSGGEVVQ 1380
Db 1321 PEVPAEVAKVLASGNWTPESPACQSQPGARRLLPDCPAAAGPPPPQAVTSGGEVVQ 1380
QY 1381 NLGTGNLSDFLVKTIPRLVROGLKTKKWNVYRGFSLGGRDPCLPQSGQELGRSVEELW 1440
Db 1381 NLGTGNLSDFLVKTIPRLVROGLKTKKWNVYRGFSLGGRDPCLPQSGQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFNNKGWSWAFVNRASNAIILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTLTAHSLDAQDSLKIFNNKGWSWAFVNRASNAIILRAHLP 1500
QY 1501 PGPARHAHSITLKNHPIMLTEQLSEALMASSVDVLVSI CVVPFAMSFPVPSFTLVLTIE 1560
Db 1501 PGPARHAHSITLKNHPIMLTEQLSEALMASSVDVLVSI CVVPFAMSFPVPSFTLVLTIE 1560
QY 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNVLPACIVVLIFLAQOYRAYVAPANLPALL 1620
Db 1561 RVTRAKHLQLMGGLSPTLYWLGNFMDMKNVLPACIVVLIFLAQOYRAYVAPANLPALL 1620
QY 1621 LLLLLYGWSITPLMYPASFFFSVPSTAYVVLTCINLFIGINGSMATFVLELFSQKLOEV 1680
Db 1589 -----KLQEV 1593
QY 1681 SRILKQVFLIIPHFCILGRGLDMVRNQAMADAFERLGRDQFQSPLRWEVGNKLLAMVIQ 1740
Db 1594 SRILKQVFLIIPHFCILGRGLDMVRNQAMADAFERLGRDQFQSPLRWEVGNKLLAMVIQ 1653
QY 1741 GPLFLFTLLQHRSQLLPQVRVSLPLIGBEDVARERERVOGATQGDVLVLRNLTJK 1800
Db 1654 GPLFLFTLLQHRSQLLPQVRVSLPLIGBEDVARERERVOGATQGDVLVLRNLTJK 1713
QY 1801 VYRGQMPAVDRCLCIGIPGECFGLLVNGAGKTSITFRMVTGDTLTSRGEAVLAGHSVAR 1860
Db 1714 VYRGQMPAVDRCLCIGIPGECFGLLVNGAGKTSITFRMVTGDTLTSRGEAVLAGHSVAR 1773
QY 1861 EPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWYAD 1920
Db 1774 EPSAAHLSMGYCPQSDAI FELLTGREHLELLARLGRVPEAQVQTAGSGLARLGLSWYAD 1833
QY 1921 RPAGTYSGNKRKLTALALVCDPAVFLDEPTTCMDPSARRFLWNSLLAVVREGRSYML 1980
Db 1834 RPAGTYSGNKRKLTALALVCDPAVFLDEPTTCMDPSARRFLWNSLLAVVREGRSYML 1893
QY 1981 TSHSMECEALCSRLAIWVNGFRCLGSPQHLKGRFAAGHTLTILRVPAARSQPAFAVAA 2040

Db 1894 TSHSMECEALCSRLAIWVNGFRCLGSPQHLKGRFAAGHTLTILRVPAARSQPAFAVAA 1953
QY 2041 BFPQSELREAHGGRRLRPQLPGGRCALARVFGELAVHGAHGVEDFVSQTMLEEVFLYF 2100
Db 1954 BFPQSELREAHGGRRLRPQLPGGRCALARVFGELAVHGAHGVEDFVSQTMLEEVFLYF 2013
QY 2101 SKDQKQDDBTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTVL 2146
Db 2014 SKDQKQDDBTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTVL 2059
RESULT 9
US-10-182-006-4
; Sequence 4, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: PCT/US01/02191
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1873
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-006-4
Query Match 84.9%; Score 9460.5; DB 15; Length 1873;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1835; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
QY 1 MAFVTQMLLLKWNFMYRRRQPVQLLVELLWPLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MAFVTQMLLLKWNFMYRRRQPVQLLVELLWPLFFILVAVRHSHPPLEHHECHFPNKP 60
QY 61 LPSAGTVPWLQGLICNVNNTCFPQITPCEBGRSLNFNDSLSVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPWLQGLICNVNNTCFPQITPCEBGRSLNFNDSLSVSRLLADARTVLGGASAH 120
QY 121 TLAGLGKLIATLRAARSTAOQPQKSPLEBPPMLDVAELLTSLLRTESTSLGALGQABPL 180
Db 121 TLAGLGKLIATLRAARSTAOQPQKSPLEBPPMLDVAELLTSLLRTESTSLGALGQABPL 180
QY 181 HSLLEAAEDLAEALLARSLVELRALLQRPRTSGPLELLSEALCSVRGSPSTVGPGLNW 240
Db 181 HSLLEAAEDLAEALLARSLVELRALLQRPRTSGPLELLSEALCSVRGSPSTVGPGLNW 240
QY 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDSDHPLSLRLLWRLLKPLILGKLLPAPD 300
Db 241 YEASDLMELVQBPESALPDSSSLSPACSELIGALDSDHPLSLRLLWRLLKPLILGKLLPAPD 300
QY 301 TPFTKRLMAQVNRITFEELTLRLDVRVWEMLGPRIFTFMDSSNVNMLQRLLOMDEGRR 360
Db 301 TPFTKRLMAQVNRITFEELTLRLDVRVWEMLGPRIFTFMDSSNVNMLQRLLOMDEGRR 360
QY 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
Db 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLEAPSEAA 420
QY 421 LVSRALQALLAHRFWAGVVFGLPEDSSDPTBHTPDGLPGHVRIRKIMDIDVVTNRKIR 480
Db 421 LVSRALQALLAHRFWAGVVFGLPEDSSDPTBHTPDGLPGHVRIRKIMDIDVVTNRKIR 480


```
QY 481 DRFDPCGADPLTDLRYVWGGFYVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPCTYDD 540
DB 481 DRFDPCGADPLTDLRYVWGGFYVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPCTYDD 540
QY 541 VFLRVLSRSLPFLTLTAWISVTLTVKAAVREKETRLDRTWAMGLSRAVLWLGWFLSCL 600
DB 541 VFLRVLSRSLPFLTLTAWISVTLTVKAAVREKETRLDRTWAMGLSRAVLWLGWFLSCL 600
QY 601 GPFLLSAAALLVVLKLGDIILPYSHPGVVFVFLAFAFVATVTSQFLLSAFFSRANLAAACG 660
DB 601 GPFLLSAAALLVVLKLGDIILPYSHPGVVFVFLAFAFVATVTSQFLLSAFFSRANLAAACG 660
QY 661 GLAYFSLYLPVYLVCVNRDRLPAGGRVAASLLSVAEFGCESLALLEEGEGQAHNVG 720
DB 661 GLAYFSLYLPVYLVCVNRDRLPAGGRVAASLLSVAEFGCESLALLEEGEGQAHNVG 720
QY 721 TRPTADVLSLAOVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPWNPFRRSYWCGPRPP 780
DB 721 TRPTADVLSLAOVSGLLLDLDAALYGLATWYLEAVCPGOYGIPEPWNPFRRSYWCGPRPP 780
QY 781 KSPAPCPTPLDPKVLVEAEAPPGLSPGVSVRSLEKRFPGSPQALRGLSLDFYQGHITAF 840
DB 781 KSPAPCPTPLDPKVLVEAEAPPGLSPGVSVRSLEKRFPGSPQALRGLSLDFYQGHITAF 840
QY 841 CHNCAGKTTTILSGLFPSPGSGAFILGHDRSSMAAIRPHLGCYQYVNVLPDMLTVDE 900
DB 841 CHNCAGKTTTILSGLFPSPGSGAFILGHDRSSMAAIRPHLGCYQYVNVLPDMLTVDE 900
QY 901 HVWFYGRKGLSAAVGPEDRLLDQVGLSKQSVQTRHLSGGMORKLSVAIAFVGGSQV 960
DB 901 HVWFYGRKGLSAAVGPEDRLLDQVGLSKQSVQTRHLSGGMORKLSVAIAFVGGSQV 960
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLAEALLDGDRVAVVAGRLCCCG 1020
DB 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLAEALLDGDRVAVVAGRLCCCG 1020
QY 1021 SPLFLRHLSGYVLTIVKARLPLTTNEKADTMEGSVDTROEKNQSGSRVCTPQLLA 1080
DB 1021 SPLFLRHLSGYVLTIVKARLPLTTNEKADTMEGSVDTROEKNQSGSRVCTPQLLA 1080
QY 1081 LVQHWVFGARLVEBPHELVLVLPYTGANDGSPATLFRELDTRLAELRLTGYSIDTSLE 1140
DB 1081 LVQHWVFGARLVEBPHELVLVLPYTGANDGSPATLFRELDTRLAELRLTGYSIDTSLE 1140
QY 1141 EIFLKVVEECAADTMEDGSCGHLCTGIAGLDVTLRLKMPPOBTALENGEPAGSAPETD 1200
DB 1141 EIFLKVVEECAADTMEDGSCGHLCTGIAGLDVTLRLKMPPOBTALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAVRVOGWALTRQOQLALLKRFLLARSRRLGCPAQIYVLPALFVGLALVPSLI 1260
DB 1201 QGSGPDVAVRVOGWALTRQOQLALLKRFLLARSRRLGCPAQIYVLPALFVGLALVPSLI 1260
QY 1261 PPGHYPALRLSPMTYGAQVSFFSDEADPGDPCARLLEALLQEAAGLEBPVQHSSHRFSA 1320
DB 1261 PPGHYPALRLSPMTYGAQVSFFSDEADPGDPCARLLEALLQEAAGLEBPVQHSSHRFSA 1320
QY 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARELLPDCPAAAGGPPPPQAVTSGSVVQ 1380
DB 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARELLPDCPAAAGGPPPPQAVTSGSVVQ 1380
QY 1381 NLTCGNISDFLVKTYPRLVROGLTKKWVNEVRVGGFSLGGRDPCLSGQELGRSVEELW 1440
DB 1381 NLTCGNISDFLVKTYPRLVROGLTKKWVNEVRVGGFSLGGRDPCLSGQELGRSVEELW 1440
QY 1441 ALLSPLPGCALDRVLKNTAWAHSILDAQDSLKIFWNNKWHSMVAFVNRASNAITLRAHLP 1500
DB 1441 ALLSPLPGCALDRVLKNTAWAHSILDAQDSLKIFWNNKWHSMVAFVNRASNAITLRAHLP 1500
QY 1501 PGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAMGFVPASFTLVLEE 1560
DB 1501 PGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAMGFVPASFTLVLEE 1560
QY 1561 RVTRAKHLQMGGLSPTLYWLGNFMDMCMNYLVPACIVVLIIFLAQOQAVAPANLPALL 1620
```

```
DB 1561 RVTRAKHLQMGGLSPTLYWLGNFMDMCMNYLVPACIVVLIIFLAQOQAVAPANLPALL 1620
QY 1621 LLLLLXGWSITPLMWYPASFPFSSVPSTAYVYVLTCLNIFIGINGSMATFVLSEFSDQKLOEV 1680
DB 1621 LLLLLXGWSITPLMWYPASFPFSSVPSTAYVYVLTCLNIFIGINGSMATFVLSEFSDQKLOEV 1680
QY 1681 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKLLAMVIQ 1740
DB 1681 SRILKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDROFQSPLRWEVVGKLLAMVIQ 1740
QY 1741 GPFLLPFLTLQLHRSOLLPOPRVRSPLLEGEEDVARERERVVQATQGDVLRNLTK 1800
DB 1741 GPFLLPFLTLQLHRSOLLPOPRVRSPLLEGEEDVARERERVVQATQGDVLRNLTK 1800
QY 1801 VYRGQMPAVDRLCGLGIPGCEGLLGVNGAGKTTSTFRMVTGDTLASRGAVALAGHSVAR 1860
DB 1801 VYRGQMPAVDRLCGLGIPGCEVSP--GVEARCRDSEMLPYCMPCSSFTEHLLCIH--- 1854
QY 1861 EPSAAHLSMG-YCPQSDAIFELL 1882
DB 1855 -----HLLLGTYCM---PIFVLL 1869
```

RESULT 10

```
US-09-995-542-2
; Sequence 2, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ulias, Iearn
; TITLE OF INVENTION: Atp-Binding Cassette Transporter-Like Molecules and
; FILE OF INVENTION: Uses thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2167
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-995-542-2
```

```
Query Match 77.3%; Score 8613.5; DB 9; Length 2167;
Best Local Similarity 76.7%; Pred. No. 0;
Matches 1678; Conservative 163; Mismatches 283; Indels 65; Gaps 11;
```

```
QY 1 MAFWTQMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPLEHCHFPNKP 60
DB 1 MAFCTQMLLLWKNTYRRRQPIQLLVELLWPLFLFFILVAVRHSHPLEHCHFPNKP 60
QY 61 LPSAGTVPMLQGLICVNNVTCFQLTGEEBGRLSNFENDSLVSRLLADARTVLGSAHR 120
DB 61 LPSAGTVPMLQGLICVNNVTCFQLTGEEBGRLSNFENDSLVSRLLADARTVLGSHIQD 120
QY 121 TLAGLGLKLIATLRAARSTAQP---QPTKQSPLEPPMLDVAELTSLRTESLGLAQ 176
DB 121 MLDALGKLIPLVRAVGGARQESDQPTSG-----SVTKLEKILQASLDPLVQ 173
QY 177 QEPHLSLLEAAEDLAQELLALRSIVELRALLQRPRTSGPLELLSEALCSVRGSSVGP 236
DB 174 QDSMRKFSDAIRDLAQELLTLPSLMELRALLRRPRGSAGSELVSEALCSTKGPSPGGL 233
QY 237 SLNWYEASDLMELVGQSPESALPDSLSLSPACSELIGALDSHPLSRLLRRLKPLILGKLL 296
DB 234 SLNWYEANQLNEFMGPEVAPALPDNSLSPACSEPFVGLDDHPVSRLLLRRLKPLILGKIL 293
QY 297 FADPTPTRKLMAQVNRFTFEELTLRLDRVREWEMLGPRIFTFMNDSSNVAMQLQLOM 356
DB 294 FADPTNTRKLMAQVNTQFEELALLRDLHELWGLVGPQIFNFMNDSTNVAMQLQLLDVGG 353
```


Qy	357	EGRQPRGGRDHMEALRSFLDPQSGGYSWQADAHADVHLVGTIGRVVTECULSLDKLEAAP	416
Db	354	TGQRQQTFRQAQKLEIAIKDFLDPGRGGYSWREAHDMGRLAGILQGMMECVSLDKLEAVP	413
Qy	417	SEALVSRALQLLAEHREFWAGVVFLLGPEDSSDPTHEPTDPLGPGHVRKIKRMDIDVVTRT	476
Db	414	SEALVSRALBELGERRLWAGIVFLSPHPLDPSELSPALSPGHLRFKIRMDIDDVTRT	473
Qy	477	NKIDRFPDGPAAADPLTDLRVYMGGFYQLQDLVERAAVRVLSGANPRAGLVLQOMPYP	536
Db	474	NKIDKFWDPGSAADPFMDLRVYMGGFYQLQDLLEQAAAVRVLGNSRGTGLVLQOMHP	533
Qy	537	YVDVFLRVLSRSLPLFTLLAWIVSVTLTVKAAVREKETRLDRTWRANGLSRAVLWLGWF	596
Db	534	YVDVFLRVLSRSLPLFTLLAWIVSVTLTVKAAVREKETRLDRTWRANGLSRAVLWLGWF	593
Qy	597	LSCGLPFLLSAALAVLVLKLGDIILPYSHPGVVFLFLAFAVATVYQFLLSAFFSRANLA	656
Db	594	LSCGLPFLVSAALAVLVLKLGNIILPYSHPPVIFLFLAFAVATVYQFLLSAFFSRANLA	653
Qy	657	AACGGLAYFSLYLPVYLVCVNRDLRPGAGRVAASLLSPVAFGFCESLALLEEGEGAQW	716
Db	654	AACGGLAYFALYLPVYLVCVNRERLHGLGLAASLLSPVAFGFCESLALLEEGDGAQW	713
Qy	717	HNVTGRTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQOYCIPEPMNPPFRSSWCG	776
Db	714	HNLTGTPAEDVFSLAQVSAFLLLDVAIVGLAWYLEAVCPQOYCIPEPMNPPFRSSWCG	773
Qy	777	PRPKSPAPCPTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQPALRGLSLDPFYOGHI	836
Db	774	PGPKSSVLAPADOPKVLVEEPGLGLVPGVSVIRGLKHPGCPQPALQGLNLDIFYEGHI	833
Qy	837	TAFLGHNGAGKTTTLLSILSGLFPSPGGSFAFLGHDRVSSMAAIRPHLGVCPQYNVLFDM	896
Db	834	TAFLGHNGAGKTTTLLSILSGLFPSPGGSASILGHDTVQTNMAAIRPHLGLICPQYNVLFDM	893
Qy	897	TVDSHVWFYGLKGLSAAVVCPEODRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIAFVG	956
Db	894	TVDSHVWFYGLKGLSAAVVCPEODRLLQDVLVSKQSVQTRHLSGGMQRKLSVAIAFVG	953
Qy	957	GSQVVIDEPTAGVDPRSRRGIMELLKRYREGRTLILSTHLDRAELLGDRVAVVAGGRL	1016
Db	954	GSRVVINDPTAGVDPRSRRGIMELLKRYREGRTLILSTHLDRAELLGDRVAVVAGGSL	1013
Qy	1017	CCGSPFLRRHLGSGYLLVLVKARLPIITNE-KADTDMEGSVDTROEKNKSGQS----	1071
Db	1014	CCGSPFLRRHLGCGYLLTVKVSQSLVTHDAKGDSE----DPRREKSDGNGRTSDT	1068
Qy	1072	-----RVCTPQLLALVOHWVPGARIVELPHELVVLVLYTCAGH	1110
Db	1069	AFTGTGSDKNQAPAGAVPITPSTARILELVQOHVPGAQVLEDPLPHELLVLVLYPYAGALD	1128
Qy	1111	GSPATLFRDLRLAELRTGCGISDTSLEEIFLKVVVEECAADTMDMDGSCGQHLCTGIA	1170
Db	1129	GSFAMVFOELDQQLLELGLTGYGISDNLNLEEIFLKVVVED--AHREGGDSRPQLHRT---	1183
Qy	1171	GLDVTLELKNPQBTALENGEPAGSA--PETDQSGGPDAGVRVOGWALTQQLQALLLKR	1228
Db	1184	---CTPQPTGPEASVLVENGELAKLVLDPOAPOGLAFNA-AQVQGWTLTCQQLRALLHKR	1239
Qy	1229	FLIARRSRGLFAQIVLPALFVGLALVFSLIVPPFGHYPALRSLPTMYGAQVFFESDAP	1288
Db	1240	FLIARRSRGLFAQIVLPALFVGLALVFSLIVPPFGYPPQLQSLPAMYGQVFFESDAP	1299
Qy	1289	GDGRARLLEALLOEAGLEBPPVQHSR-----FSAPEVPAEVAKVLASGNWT	1337
Db	1300	GDPRMKLLEALLGEAGLQESPMQDKDARGSECTHSLACFTYTPVEVPPDVASILASGNWT	1359
Qy	1338	PESGPSACQSQPCARLLPDCPAAAGGPPPPQAVTSGSEVQNLTCGNLSDFLVKTYPR	1397
Db	1360	PESGPSACQSQPCARLLPDCPAGAGGPPPPQAVAGLGSEVQNLTCGNVSDFLVKTPS	1419

Qy		LVRGLTKKWKVNEVRYGGFSLGRDPGLSPGQELGRSVBELWALLSPLPGGALDRLVKN	1457
Db		LVRRGLTKKKWDVEVRYGGFSLGRDPDLPTGHGVVRTLAEIRALLSPQGNALDRILNN	1479
Qy		LTAWAHLSLDQDSIKTWFNNKGWHSVMVAFYVNRSANILLRAHLPPGPAPARHAHSITTLNHPL	1517
Db		LTOVALGDARNLSUKWFNNKGHWAMVAFYVNRANNGLLHALLPSPGVPVRHAHSITTLNHPL	1539
Qy		NLTKEQLSEAAUMASSVDVLVSICVVFAMSFPASPFTVLIEBRVTTRAKHLQMLGGLSPT	1577
Db		NLTKEQLSEATLIASSVDVLVSICVVFAMSFPASFTVLIEBRITTRAKHLQVLGSLPQT	1599
Qy		LYMIGNFLDMCNVLPACIVVLFIAPAORAIVA PANIPALLLLLLLLXGWSITPLMYPA	1637
Db		LYMLGNFLDMCNVLAVCIIWPIFIAPAORAIVA PENIPALLLLLLLLYGWSITPLMYPA	1659
Qy		SFFFSVPSTAYVVLTCINLFIGINGSMAFPVELFSDOKLOEVSRTLKVFLIFPHFCUG	1697
Db		SFFFSVPSTAYVVLTCINLFIGINSMAFPVELSDQNLOEVSRTLKVFLIFPHFCUG	1719
Qy		RGLIDMVNRQAMADAPERLGDRQFQSPLRWEVVGNKLAWIQCPLFTLLTLLQHRSOL	1757
Db		RGLIDMVNRQAMADAPERLGDKQFQSPLRWDIIGKNLLAMMAQGPIFLTITLLQHRNL	1779
Qy		LPQDRVSRPLPLEGEEDVARERERVVOGATOGDVLVLNRLTKVYRGORMPADVRLCLGI	1817
Db		LPQSKPRLLPPLGEEDVAERERTVKGATOGDVLVLRDLTKVYRGORNPAVDRLCLGI	1839
Qy		PPECFGLLGVNCAGKTSTRMWTDGTLASRGAENVLAGHSVAREPSAHLSMGYCPOSDA	1877
Db		PPECFGLLGVNCAGKTSTRMWTDGTLPSSGBAVLAGHNVAQERSAAHRSMGYCPOSDA	1899
Qy		IFELLTGREHLELLARLVPEAQVAQTAGSGLARGLGSWYADRPACTYSGGNKRKATA	1937
Db		IFDLLTGREHLELFAIRLVPEAQVAQTALSGLVRLGPLSYADRPACTYSGGNKRKATA	1959
Qy		LALVGDPDVVFLDEPTTGMDPSARRFLMNSLLIAVREGRSVMLTSHSMECEBALCSRLAI	1997
Db		LALVGDPDPAVVFLDEPTTGMDPSARRFLMNSLLISVVREGRSVVLTSMSMECEBALCTRLAI	2019
Qy		MVNGRFRCILGSPHLLKGRFAAGHTLLRLVPAARSOPAAAFAVAEPGSELREAHGGLRPF	2057
Db		MVNGRFRCILGSSOHLKGRFGAGHTLLRLVPPQPEPAIFIPIITPPGAELREVHGSRILRF	2079
Qy		QLPPGGRCALARYFGELAVHGAHEGVDEFVSQTMLEEVEFLYFSKDQKHDEDTTEEQEAG	2117
Db		QLPPGGRCTLTRYFRELAQAAGRAHVDEFSVQGTILLEEVFLYFSKDQGEEBESSRO-EAB	2138
Qy		VGVDPAPGLQHPKRVSQFLLDDPSTAETVL	2146
Db		EEEVSKPGROHPKRVSRFLEDSPSSVBTFMI	2167

```

RESULT 11
US-09-995-542-3
; Sequence 3, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2121
; TYPE: PRT
; ORGANISM: Mus musculus

```

US-09-995-542-3

Query Match		75.3%; Score 8389.5; DB 9; Length 2121;
Best Local Similarity		76.3%; Pred. No. 0;
Matches 1636; Conservative 161; Mismatches 281; Indels 65; Gaps 11;		
Qy	47	PPLHHECHFPNKLPSAGTVPWLOGLICVNNVTCFQLPITGPEBPGRLSNFNDLSVRL 106
Db	1	PPLHHECHFPNKLPSAGTVPWLOGLVNVCNNSCFQHPPTGKPGVGLSKDLSIRLL 60
Qy	107	ADARTVLGGASARTLAGLKIATLRAARSTAPQ- - - OPTKQS PLEPPMLDVAEILTS 162
Db	61	ADRTVLCGHSIQMDLALGKLIIPVLRVAGGARQESDQTSQG- - - - - SVTKLEK 113
Qy	163	LLRTESLGLAQOAEPLHLSLEAAEDLAQELLALRLSLVELRALLQRPRTGSGPLELSE 222
Db	114	ILQASLDPLVGLQAQDSMRKFSDAIRDLAQELLTLPSLMELRALLRRPRGSAGSLELVSE 173
Qy	223	ALCSVRGPGSSTVGPSLWNYEASDLMELVQBPESALPDSSLSPACSELIGALDHPISRL 282
Db	174	ALCSTKGPSSPGGLSLWNYEANQNEFMGPEVAPALPDNLSLPACSEFVGTLDHPVSRL 233
Qy	283	LWRLKPLILGKILFAPDPTTRKLMQVNRTPFEELTLRDVREVMGLGPRIETFWNDS 342
Db	234	LWRLKPLILGKILFAPDPTNTRKLMQVNOTFEELALLRDLHELWGLGPQIFENFWNDS 293
Qy	343	SNVAMQLQOMODEGRRQPPGGRDHMEALRSFLDFSGSGYSQDADHADVGLVGTGR 402
Db	294	TNVAMQLRLDVGCTGQOQTFRACKLEAIKDFLDPSPRGYSWREAHADMGRLAGILQ 353
Qy	403	VTECLSDKLEAAPESEALVSRLQALLAEHRFWAGVVPFLGPESSDDTEHPTDLPGRHV 462
Db	354	NMECVSLDKLEAVPESEALVSRLLELGERLMAGIVFLSPEHPPLDSPSELSPALSPGHL 413
Qy	463	RIKRMDIDVTRNKIRDRFMDGPAADPTDLRYVMGPFVYLQDLVERAAVRVLSCAN 522
Db	414	RFKRMDIDVTRNKIRDKFWDGPGSADPMDLRYVMGPFVYLQDLLEQAAVRVLGGN 473
Qy	523	PRAGLYLQOMPYPCYDDVFLRVLRSPLPLTLTAWIYSVTLTKAVVREKETRLRDTMR 582
Db	474	SRTGLYLQOMPHPCYDDVFLRVLRSPLPLTLTAWIYSVALTKAVVREKETRLRETMR 533
Qy	583	AMGLSRVILWGLFSLGCLPPLLSAALLVVLKGDILPYSHPGVFLFLAAFAVATVQ 642
Db	534	AMGLSRVILWGLFSLGCLPPLVSAALLVVLKGNILPYSHPVVIFFLAAFAVATVQ 593
Qy	643	SFLLSAFSSRANLAAACGLAYFSLXLYPYVLCVAMRDLRPAAGRVAASLLSPVAFPGCE 702
Db	594	SFLLSAFSSRANLAAACGLAYFALYPLYVLCVAMRERLHGLLAAASLLSPVAFPGCE 653
Qy	703	SLALLEBQGEAOWHNVGTRPTADVFLSAQVSGLLLLDAALYGLATWYLEAVCPQYQYIP 762
Db	654	SLALLEBQGEAOWHNLGTGPAEDVFLSAQVSAFLLLDAVIYGLALWYLEAVCPQYQYIP 713
Qy	763	EPWNFPFRSRYWCGRPPKSPACPCTPLDPKVLVEAPPGLSPGVSYSRLEKRPSPGP 822
Db	714	EPWNFPFRSRYWCGRPPKSSVLTAPAPQDPKVLVEEPPPLGLVPGVSRGLKGRPGCPQ 773
Qy	823	ALRGLSLDFYOGHITAFILGHNGAKTTLTSLGFLFPSPGGSATILGHDRVSSMAAIRPH 882
Db	774	ALQGLNLDYFEGHITAFILGHNGAKTTLTSLGFLFPSPGGSASILGHVQTNMAAIRPH 833
Qy	883	LGVCPOYNVLPDMULTVDEHVMFYGRLLKGLSAANVGPBQDRLLQDVLVSKOSVOTRHLSG 942
Db	834	LGICPOYNVLPDMULTVEHVMFYGRLLKGLSAAAMGPERELIRDVGLTLKRDQTRHLSG 893
Qy	943	GMORKLVAIAFVGGSQWILDETAGVDPASRGRIWELLILKYREGTTLILSTHLLDEAE 1002
Db	894	GMORKLVAIAFVGGSRVIMDEBTAGVDPASRRGIWELLILKYREGTTLILSTHLLDEAE 953
Qy	1003	LLGDRVAVAGRLCCCGSPFLFRHLGSGYYLTLVKARPLTATNE- KADTDMEGSVDTTR 1061
Db	954	LLGDRVANVAGSLCCCGSPFLFRHLGCGYYLTLVKSSQSLVTHDAKGDS- - - - - DPR 1008

Qy	1062	QEKKNSSQGS- - - - - RVGTPQALLVQHWVPGARLVEELP 1096
Db	1009	REKSDNGRSTDTAFTRGTSDKSNQAPAGVPIPTSTPARIILELVQOHPGQALVDELPL 1068
Qy	1097	HELVLVLPYTGADGSPATLPRLELDTRLAELRLTGYSIDTSLLEEILFKVVECAADTM 1156
Db	1069	HELLVLFPYAGALDGSFAMVQELDQOELLGLTGYSIDTNLEEILFKVED- - AUREG 1126
Qy	1157	EDGSCGHLCTGAGLDVTLRLKMPQOETALENGEPAGSA- - PETDQSGSPDAVGRVQGW 1214
Db	1127	GDSPQLHLRT- - - - - CTQPPTGPEASVLENGELAKLVLDPOAQGLAPNA- AQVQGW 1179
Qy	1215	ALTRQOQALLKRLFLARRSRGLFAQIVLPALFVGLALVFLSVLPPFGHYPALRUSPT 1274
Db	1180	TLTCQQLRALLHRLFLARRSRRLFAQVVLVLPALFVGLALFFSLIVPPFGQYPLQLSPA 1239
Qy	1275	MYGNOVSFFSEDAPODGRARLLEALLQEALEAGLEPPVQHSR- - - - - FSAPEV 1323
Db	1240	MYGQVSFFSEDAPODNRMKLLEALLGEAGLEQEPSMDKARGSECTHSLACYFTVPEV 1299
Qy	1324	PAEVAKVLASGNWTPSPSPACQCSQPGARRLLPDCPAAAGGPPPPQAVTSGSEVVQNLT 1383
Db	1300	PPDVASILASGNWTPSPSPACQCSQPGARRLLPDCPAGAGGPPPPQAVAGLGEVVQNLT 1359
Qy	1384	GRNLSDFLVKTYPRLVROGLKTKKWNVEVYGGFSLGGRDPGLPSGOELGRSVEELWALL 1443
Db	1360	GRNVSDFLVKTYPSLVRRGLKTKKWDENVYGGFSLGGRDPDLPTGHEVVRTLAEIRALL 1419
Qy	1444	SPLPGGALDRVLKNTLTAWAHSLDAOSLKTWFKNGHSMVAFVNRASNALILRAHLPGP 1503
Db	1420	SPOFGNALDRILNLTQWALGLDARNSLKTWFKNGHSMVAFVNRANNGLHALLPSGP 1479
Qy	1504	ARRAHSITTLNHPNLNTKEQLSEAAALMASSVDLVLSI CVVVFAMS FVPASFTLVLIBERT 1563
Db	1480	VRHAHSITTLNHPNLNTKEQLSEATL IASSVDLVLSI CVVVFAMS FVPASFTLVLIBERT 1539
Qy	1564	RAXHLQMLGSLSTPLWGNFLMDMCMYLPACIVVLI FLAFOORAYVAPANLPALLLLL 1623
Db	1540	RAXHLQVSLGPOTLWYLGFLMDMCMYLVAVCVVIFL AFOORAYVAPENLPALLLLL 1599
Qy	1624	LIXGWSITPLMYPASFPFSPVSTAYVVLTCINLFIGINGSMATFVLBLEFSDQLQEVSR 1683
Db	1600	LLYGWSITPLMYPASFPFSPVSTAYVVLTCINLFIGINSSMATFVLELLSDONLQEVSR 1659
Qy	1684	LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRMEVVYKKNLLAMVIOGPL 1743
Db	1660	LKQVFLIFPHFCLGRGLIDMVRNQAMADAFERLGDQFQSPLRWDIIGKNLLAMMAQGPL 1719
Qy	1744	FLITLTLLOHRSOLLPOQPRVRSIPLIGEEDVARERERVQATQGDVLVRLNLTKVYR 1803
Db	1720	FLITLTLLOHNRLLPOSKPRLLPPLIGEEDVARERERVTKGATQGDVLVRLDRLTKVYR 1779
Qy	1804	GORMPADVRLCIGIPGCECFGLLVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAREPS 1863
Db	1780	QORNPADVRLCIGIPGCECFGLLVNGAGKTSITFRMVTGDTLPSSGEAVLAGHVAQERS 1839
Qy	1864	AAHLSMGYCPQSDAIIFELLTGREHLELLARLRGYPEAQVATAGSGLARLGLSWYADRP 1923
Db	1840	AAHLSMGYCPQSDAIPDLITGREHLELLARLRGYPEAQVATALSGLVRLGLPSYADRP 1899
Qy	1924	GTYSGGNKRKLATALALVGDPAVVFLEPTTGMDDPSARRFLNLSLLAVREGRSVMLTSH 1983
Db	1900	GTYSGGNKRKLATALALVGDPAVVFLEPTTGMDDPSARRFLNLSLLSVREGRSVMLTSH 1959
Qy	1984	SMECEALCSRLATMNGRFRCLGSPHOLKGRFAAGHTLTLRVPAAASQPAAPFAVAEFP 2043
Db	1960	SMECEALCTRILAMNGRFRCLGSSOHLKGRFCAGHTLTLRVPPDQPEPAIAFIRITFP 2019
Qy	2044	GSELREAHGRLRFLQPPGRCALARVFGELAVHGAHGVEDFSVQTMLEEVFLYFSKD 2103
Db	2020	GAELEHVHGSRLRFLQPPGRCCTLRVRELAAGRAHGVEDFSVQTMLEEVFLYFSKD 2079

Qy 2104 QKQEDTEQKEAGVDPAPGLOHPKRVSQFLDDPSTAETVL 2146
Db 2080 QGEESSESRQ-EABEEVSKPGRQHPKRVSRFLEDPSSVEITMI 2121

RESULT 12

US-09-995-542-8
; Sequence 8, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; FILE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; PRIOR APPLICATION DATE: 2001-11-28
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1550
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-995-542-8

Query Match 68.2%; Score 7596; DB 9; Length 1550;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 167 ESLGALQAQEPHLSLEAAEDLAQELLALRSVLRAQLRPRGTSGPLELSEALCS 226
Db 29 ESLGALQAQEPHLSLEAAEDLAQELLALRSVLRAQLRPRGTSGPLELSEALCS 88

Qy 227 VRGSSSTVGPSINWYEASDLMELVQEPESALPDSSLSPACESLIGALDHSPLSRLLWR 286
Db 89 VRGSSSTVGPSINWYEASDLMELVQEPESALPDSSLSPACESLIGALDHSPLSRLLWR 148

Qy 287 LKPLILGKLLFAPDTPFRKMAQVNRTPFELTLRDVREVMELGPRIFTFMDSSNVA 346
Db 149 LKPLILGKLLFAPDTPFRKMAQVNRTPFELTLRDVREVMELGPRIFTFMDSSNVA 208

Qy 347 MLQRLQMDGRRPRGGRDHMEALRSFLDPGSGYSWODAHADVGLVGTLCRVTEC 406
Db 209 MLQRLQMDGRRPRGGRDHMEALRSFLDPGSGYSWODAHADVGLVGTLCRVTEC 268

Qy 407 LSLDKLEAAPSEALVSRALQLLAHREFWAGVVFGLGPDSDPTHEPTDLPDGPVHRIKI 466
Db 269 LSLDKLEAAPSEALVSRALQLLAHREFWAGVVFGLGPDSDPTHEPTDLPDGPVHRIKI 328

Qy 467 RMDIDVTRTKIRDRFWDGPAADPLTDLYRWGGFVYLODLVERAAVRVLSGANPRAG 526
Db 329 RMDIDVTRTKIRDRFWDGPAADPLTDLYRWGGFVYLODLVERAAVRVLSGANPRAG 388

Qy 527 LYLOQMPYPCVDDVFLRAVLSRSLPFLTLTAWISVTLTVKAVVREKTRLDTRAMGL 586
Db 389 LYLOQMPYPCVDDVFLRAVLSRSLPFLTLTAWISVTLTVKAVVREKTRLDTRAMGL 448

Qy 587 SRAVLWLGWFLSCLPGLPFLLSAALLVLVLKLDGILPYSHPGVVFLLAFAVATVTSQFLL 646
Db 449 SRAVLWLGWFLSCLPGLPFLLSAALLVLVLKLDGILPYSHPGVVFLLAFAVATVTSQFLL 508

Qy 647 SAFFSRANLAACGLAYFSLYPLVLCVAVNRDLRACGRVAASLLSPVAFGFCESIAL 706
Db 509 SAFFSRANLAACGLAYFSLYPLVLCVAVNRDLRACGRVAASLLSPVAFGFCESIAL 568

Qy 707 LEEQEGQAWNVGTRPTADVFSLAQVSGLLLDLAALYGLATWYLEAVCPQYGIPEPWN 766
Db 569 LEEQEGQAWNVGTRPTADVFSLAQVSGLLLDLAALYGLATWYLEAVCPQYGIPEPWN 628

Qy 767 FPFRRSYWCGPRPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQALRG 826

Db 629 FPFRRSYWCGPRPKSPAPCTPLDPKVLVEEAPGLSPGVSVRSLEKRPFGSPQALRG 688
Qy 827 LSLDFYQGHITAFILGHNGAGKTTTILSLGILFPPSGGSAPILGHDVRSMAAIRPHLGYC 886
Db 689 LSLDFYQGHITAFILGHNGAGKTTTILSLGILFPPSGGSAPILGHDVRSMAAIRPHLGYC 748

Qy 887 PQYNVLFDMLTVDHFWFYGRLKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMOR 946
Db 749 PQYNVLFDMLTVDHFWFYGRLKGLSAAVVGPEQDRLLQDVGLVSKQSVQTRHLSGGMOR 808

Qy 947 KLSVAIAFVGSGVQVILDEPTAGVDPPASRRGIWELLKYRGRTILSTHHLDEALDGD 1006
Db 809 KLSVAIAFVGSGVQVILDEPTAGVDPPASRRGIWELLKYRGRTILSTHHLDEALDGD 868

Qy 1007 RVAVVAGGRLCCGSPPLRRLHLSGYVLTAVKARLPLTTNEKADTMEGSDVTROEKN 1066
Db 869 RVAVVAGGRLCCGSPPLRRLHLSGYVLTAVKARLPLTTNEKADTMEGSDVTROEKN 928

Qy 1067 GSQGSRVGTPQLLALVQHWVPCARLVEPHELVLVLPYTGADHGSFATLFRDLTRLAE 1126
Db 929 GSQGSRVGTPQLLALVQHWVPCARLVEPHELVLVLPYTGADHGSFATLFRDLTRLAE 988

Qy 1127 LRLTGYSIDTSLBEIFLKVVVECAADTDMEDGSCQHLCCTGIAGLDTVLRKMPQETA 1186
Db 989 LRLTGYSIDTSLBEIFLKVVVECAADTDMEDGSCQHLCCTGIAGLDTVLRKMPQETA 1048

Qy 1187 LENGEPAQSAPETDQSGPDAGVQGWALTRQQLQALLKRLFLARSRRLFAQIVLP 1246
Db 1049 LENGEPAQSAPETDQSGPDAGVQGWALTRQQLQALLKRLFLARSRRLFAQIVLP 1108

Qy 1247 ALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVYFFSEDAFGDPGRARLLLEALQEAGL 1306
Db 1109 ALFVGLALVFSLIIVPPFGHYPALRLSPMYGAQVYFFSEDAFGDPGRARLLLEALQEAGL 1168

Qy 1307 EEPVQVSHSRFSAPEVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1366
Db 1169 EEPVQVSHSRFSAPEVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1228

Qy 1367 PPQAVTGSVGVQNLTCRNLSDFLVKTYPRLVQGLTKKWNVNEVYGGFSLGGRDPL 1426
Db 1229 PPQAVTGSVGVQNLTCRNLSDFLVKTYPRLVQGLTKKWNVNEVYGGFSLGGRDPL 1288

Qy 1427 PSQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLDQDSLKIFWNNKGWSMVA 1486
Db 1289 PSQELGRSVEELWALLSPLPGGALDRVLKNTAWAHSLDQDSLKIFWNNKGWSMVA 1348

Qy 1487 VNRSNAILRAHLPPGPARHAHSITTLNHPNLTKQQLSEALMASSVDVLVSVICVVFAM 1546
Db 1349 VNRSNAILRAHLPPGPARHAHSITTLNHPNLTKQQLSEALMASSVDVLVSVICVVFAM 1408

Qy 1547 SFVPASPTLVLEERVTRAKHLQLMGGLSPTLYLGNFLWDMCNVLPACIVVLIPLAFQ 1606
Db 1409 SFVPASPTLVLEERVTRAKHLQLMGGLSPTLYLGNFLWDMCNVLPACIVVLIPLAFQ 1468

Qy 1607 QRAVVAANLPAALLLLLLYGSITPL 1633
Db 1469 QRAVVAANLPAALLLLLLYGRQAVPV 1495

RESULT 13

US-10-313-641-9
; Sequence 9, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P02351US2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06

```
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
; US-10-313-641-9

Query Match
Best local Similarity 51.8%; Score 5768.5; DB 14; Length 2261;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;

Qy 1 MAFWTQLMLLWKNFYRRQVQLLYVLLWPLFLPILVAVRHSHPPLSHHECHRNKP 60
Db 1 MACWPQLRLLLWKNLTFRRTQTCQLLEAVWPLFIFLILISVRLSYPPYQHECHFNKA 60
Qy 61 LPSAGTVPWLQGLICNVNNTCTFPQLTGEEBPGRLSNFNSLVSRLLDARTVLGASAH 120
Db 61 MPSAGTLPWQGLICNANNPCFRYPPTGEPAGVVGNGFNKSVARLPSDARRLLYSQDT 120
Qy 121 TLAGLGLIATLRAARSTAQ-----PQTKOSPLEPPML-- 154
Db 121 SMKDMKVLRTLOQIKSSSNLKLQDFLVNFTSGFLYHNLSPKSTVDKMLRADVILH 180
Qy 155 -----DVAEL-----LTSLL 164
Db 181 KVFLQGVQLHLTSLCNGSKSEMTQLGDQVSELGCLPREKLAARVLRNSMDILXPIL 240
Qy 165 RT--ESLGLAQOAEPLHLSLEAAEDLAQELALRSLVELR---ALLQRPRTSGPLEL 219
Db 241 RTLNSTSPFSPKELAEATKTLHLSLGTLAQELFPMRSWSDMRQBWMLFTNVNWSSTQI 300
Qy 220 ---LSEALCVRGSPSTVGSPLNWEASDLMELVG----QEPESALPDSSLSPACSELIG 272
Db 301 YQAVSRIVCGHPEGGLKIKSLNMYEDNNYKALFGGNGTEDEATFYDNSTTTPYCNLAK 360
Qy 273 ALDSHPRLRLWRLKPLILGLKLFADPTTPTRKLMQVNRTPBELTLDRVREWBMLG 332
Db 361 NLESSPLSRIIWKALKPLLVGKILYTPDTATROVMAEVNKTFOELAVFHDLEGWELSL 420
Qy 333 PRIFTMNDSSNAMLQRLLOMQDEGR-RQRPQGRD-----RHEALRSFLDP 379
Db 421 PKIWMENSQEMDLVRMLDSDRNDHFWEQQLDGLDWTQADI VAFLLAKHPEDVQS---- 476
Qy 380 GSGG--YSWQDAHADVGHVGLTGRVTECLSLDKLEAAPSEALVSRALQLLAHRFWAG 437
Db 477 -SNGSVYTWREAFNETNQARTISRFMECVNLNKLPIATEVWLNKSMELLDSEKFWAG 535
Qy 438 VFLGPDSDSPTBHTPOLGPGHVRIRKIMDIDVTRTKIRDFWDPGPAADPLTDLR 497
Db 536 IVFTGTPGSTELPH-----HVYKIRMDINDVERTNKIKDGYWDPGRADPPEDMR 587
Qy 498 YVWGFGVYLDQVLRRAVVRVLSGANBRAGLYLOQMPYPCYVDDVFLRSLRSLPLFLTLA 557
Db 588 YVWGFGFAYLDQVVRQAIIRVLVTGKKTGVTYMQMPYPCYVDDIFLVRMSRSMPLFWTLA 647
Qy 558 WIYSVTITVKAIVREKETRLDTRMAGLSRAVLWGLFSLCLGPFLLSALLLVLVLKLG 617
Db 648 WIYSVAVIIKIGIVYEKEARLKTNRIMGLNDSILWFSWFTSSLIPLVSVAGLLVVLKLG 707
Qy 618 DILPSHPGVVFLFALAFAVATVQSFLLSAFFSRANLAAACGLAFSLYPLPVLCVAV 677
Db 708 NLLPSDPSVVFVLSVFAVVTIILQCFILSTLFSRANLAAACGIIIFTLYLPLVLCVAV 767
Qy 678 RDRLPACGRVAASLLSPVAFGCGESLALIEEQEGEQAGHNVGTRPT-ADVLSLAQVSLG 736
Db 768 QDYVGFILKI PASLLSPVAFGCGEYFALFEQOIGVQWDLNLPSPVEEDGFNLTTTSM 827
Qy 737 LLLDAALYGLATWLEAVCPQOQYIPEPWNPPFRSYWCGRPKSPACPTPLD-PKVL 795

828 MLFDTLYGVMTWYIEAVFPQGIIPRPWTPTCKTSWFGESDEKSHPSGNQKRISEIC 887
796 VEEAPPGLSPGVSRSLEKRFPGSPQALRGLSLDFYQGHITAFGLHNGAGKTTTILS 855
888 MEEEPHLKLGVSQNLVKYRDGMKVAVDGLALNFVEGQITSLFGLHNGAGKTTMSILT 947
856 GLFPSPGSAFILGHVRSMAAIRPHLGVCPQYNVLFMLTVDEHVVFFGRKLGLSAV 915
948 GLFPPTSTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHWFARLKLSEKH 1007
916 VGPEQRLLLQDVL-VSKQSVQTRHLSGGMQRKLSVAIAFVGSGSVVILDEPTAGVDPAS 974
1008 VKAEMEQMALVDPLSPKSKTSQSLSGMQRKLSVALAFVGSKVILDEPTAGVDPS 1067
975 RRGIWELLKYREGRTLIILSTHLDLAEALGDRVAVVAGGRCLCCGSPFLRRLHLSGY 1034
1068 RRGIWELLKYRGRTIILSTHMDLADVLGDRIAIISHGKLCCVSSSLFLKNQLGTGY 1127
1035 LTLVKARLPLTTN-----EKADTDMGSDVTROEKNKGSGSGRVGTFPQLLALV 1082
1128 LTLVKKDVESLSSCRNSSSTVSYLKEDSVSSSDAGLGSDESDTLTIDVSAISNLI 1187
1083 QHVPVPGARLVEELPHELVLVLPYTGADGSAFATLRELDTRLAELRLTGYGISDTSLEE 1142
1188 RKHVSARLVDELIGHELTYYLFPYAAKEGAFVLEFHEIDRLDGLGISSGISETTLEE 1247
1143 FLKVBECAADTDMEDGSCQHLCCTGJAGLDVTLRLKMPQPTALE--NGEPAGSAPETD 1200
1248 FLKVAESGVDAETSDGTLPARNRRAFG-DKQSLRPFTEDDAADPNDSDIDPESRETD 1306
1201 QSGSPDAVG--RVQGWALTRQLOALLKRFILARRSRRLGFAQIVLPALPVGLALVPSL 1258
1307 LLSGMDGKGYQYKWKLTQQQFVALLWKRLIARRSRKGFQIVLPFAVFCIALVPSL 1366
1259 IVPFGHYPALRLSPMYGAQVSFFSEDAFGDPRARLLEALLQEAQ-----LEE 1308
1367 IVPFGKYPSELPQWYNEQYTFVNDADPTGTLELLNALTDPGFGTRCMEGNIPD 1426
1309 PPVQSHSRESAPEVAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGPPP 1368
1427 TPCQAGEEETAPVPTQIMDLFQNGNWTQNPSPACQSSDKIKMLPVPCCGAGLPP 1486
1369 PQAVTSGEVVQNLTGNSLDFLVKTYPRLVROGLTKKMWNEVRYGGSIG-GRDPLP 1427
1487 PQKQNTADLQDLTGRNISDYLVKTYVQIIAKSLKNKIWNNEFRYGGFSLGVSNTQALL 1546
1428 SGQLGRSVBELWALLSPLPGGALDVLKMLTAWHSLDAQDSLKIFNKNKGWSHMAFV 1487
1547 PSQEVNDAIKQMKKHLKXADSSADRFNLGLRPMTCGLDTKNVKNVFNKNKGWHAISFL 1606
1488 NRASNAILRAHLPPGPARHAHSITTLNHLPLNLTKEQLEAALMASSVDVLVSVICVFPMS 1547
1607 NVINNALIRANLQKGENPSHYGITAFNHPNLNLTQQLSEVALMTTSVDVLVSVICVFPMS 1666
1548 FVPASFTLVILBERVTRAKHLQMLGSLPTLYLGNFLWDMCNVLPACIVLFLFAQQ 1607
1667 FVPASFVFLIQBRVSKAKHLQFISGVKPYIYLSNFWDMCNVVPATLIIIFICFQQ 1726
1608 RAYVAPANLALLLLLVGWSITPLMYPASFFSPSPSTAYVVLTCINLFIGINGSMATF 1667
1727 KSVSVSNLPLVALLLLLVGWSITPLMYPASFFKIPSTAYVVLTSVNLFIGINGSVATF 1786
1668 VLELFSQDKLOEVSRIKQVFLIPPHFCLGRGLIDMVNRQAMADAFERLGRDQFQSPLRW 1727
1787 VLELFTDKLNNIDILKSVFLIPPHFCLGRGLIDMVNRQAMADALERFGRNRRVSPLSW 1846
1728 EVVGKILLAMVIOGPLELFTLLLOHRSQLLPQPRVRSPLLGEEDDVARERERVQGA 1787
1847 DLVGRNLFAMAVGVVFFLTIVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQIIDGG 1906
1788 TQGVDLVRLRLTKVYRGORMPAVDRLCLGIPPGCEGFLGLGVNGAGKTTSTFRWVTGDTLAS 1847
```

Db 1907 GONDILEIKELTIYIRKRPKPAVDRI CVGIPPECFGLLVNGAGKSSTFKMLTDTTVT 1966
Qy 1848 REEAVLAGHSVAREPSAAHLSWCYCQSDAIPELLTGREHLLEAARLRCVPEAQVATAG 1907
Db 1967 RGDAFLNKNISLNSNHEVQNMGCYQPDPAITELLTGREHVEFFALLRGVPEKEVGKGE 2026
Qy 1908 SGLARIGLSWYADRPAQTYSGGNKRKLATALVGDPAVVFDEPTTGMDSARRPLMNS 1967
Db 2027 WAIRKLGVLKYGEKVAAGNSGCKRKLSTAMALIGGPPVVFDEPTTGMDSARRPLMNC 2086
Qy 1968 LLAVVREGSVMLTSHSMECEALCSRLAIVNGRPRCLGSOHLKPRGAAGHTTLRVP 2027
Db 2087 ALSVVKESGVVLTSHSMECEALCTMAIMVNGRPRCLGSOHLKPRFGDGTIVVRIA 2146
Qy 2028 AARS--OPAAAFAAEFFPGSELREAGGRLRFQLPGGRCALARVFGELAVHGAHGVED 2085
Db 2147 GSNPDLKPVDFFGLAFPGSVLKEHRNWLQQL--PSSLSUARIFSILSQSKRLHLIED 2205
Qy 2086 FVSQTMLEEVLYFSKQKDBE---DTEEOKEAGVGVDPAPGLQHPKRVSOFLDDPSTA 2142
Db 2206 YSVSQTLDDQVFNFAKQSDDDHLKLSLHKNQTV--VDVAV-----LTSFLQDEKVK 2257
Qy 2143 ET 2144
Db 2258 ES 2259
RESULT 14
US-10-313-641-10
; Sequence 10, Application US/10313641
; Publication No. US20030162758A1
; GENERAL INFORMATION:
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P023151US2
; CURRENT APPLICATION NUMBER: US/10/313,641
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
US-10-313-641-10
Query Match 51.8%; Score 5768.5; DB 14; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;
Qy 1 MAFWTQLMLLWKNFMYRRQPVQLLVLLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60
Db 1 MACWPQLRLLWKNLTFRRQTCQLLEAVNPLFLFILISVRLSYPPPYEQHECHFPNKA 60
Qy 61 LPSAGTVFWLQGLICNVNNTCPQLTPGPEPRLSNFNDLSVRLADARTVLGGASAH 120
Db 61 MPSAGTLFWQGLICNVNNTCPYTPGEPGVGNFNKSVARLFSRALLLYSQDXT 120
Qy 131 TLAGLGLKLIATLARAARSTAQ-----POPTKQSPLEPPML-- 154
Db 121 SMKWRKVLRTLQIKKSSNLKQDLVDNFTSGFLYHNLSPKSTVDKMLRADVLH 180
Qy 155 -----DVNEL-----LTSLL 164
Db 181 KVFLQGYQLHLTSLCNGSKSEMIQLGQGEVSELGCLPREKLAARVLRNSMDILKPI 240

Qy 165 RT--BSLGLALGQAQBPPLHSLLEAAEDLAQELLALRSLVELR---ALLQRPRTSGPLEL 219
Db 241 RTLNTSTSPPPSKELAEATKTLHSLGTLAQELFSPWRSWDMRQEVWFLTNVNSSSTQI 300
Qy 220 ---LSEALCSVRGSPSTVGPSSLNWYEASDMLVLG---QEPESALPDSLSLSPACSELIG 272
Db 301 YQAVSRI VCGHPGEGGLKI KSLNWYEDNNYKALFGNGNGTEEDAETPDNSTTYPYCNLAK 360
Qy 273 ALDSPLRLKRRRLKPLILGKLLFAPDTPPTFKLMAQVNRTEELTLRLDRVREVMELG 332
Db 361 NLESSPLSRIIWKALPLLVGKILYTPDTPATQWMAEVNKTQFELAVFHELGMMEEELS 420
Qy 333 PRIFTFMDSNVAMQLRQLQWQDEGR--RQRPGGGRD-----HMEALRSFLDP 379
Db 421 PKIWTFMENSOEMDLVRMLLDSRDNDHFWEQQLDGLDWTQADIVAFVLAHPEDVQS--- 476
Qy 380 GSGG--YSQDADHADVHLVGLTRVTECLSDKLEAPSEALVSRAQLQLAHRFWAG 437
Db 477 -SNGSVYTWREAFNETNQAIRTISRMECVNLKLEPIATEVNLINKSMELDERKFWAG 535
Qy 438 VVFLGPESSDPTHTPDLPGLGHVRIKTRMDIDVVTNRKTRDRFWDGCPAADPLTLDR 497
Db 536 IVFTGITPGSIELPH-----HVYKIRMDIDNVERTNKIKDGYWDGPRADPEDMR 587
Qy 498 YWVGFFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPYCYVDDVFLRVLRSRSLPLTLA 557
Db 588 YWVGGFAYLQDVVEQAIIRVLGTGTEKTVGYMQOMPYPYCYVDDIIFLRVMSRSMPLFMTLA 647
Qy 558 WIYSVTLTVKAVVREKETRLDRTRAMGLSRAVLWGLWPLSLCLGPFPLISAALLVLKLG 617
Db 648 WIYSVAVIIGIVYEKEARLKEKTRIMGLDNLWFSWFISSILPLLVSAGLVLVILKLG 707
Qy 618 DILPSHGVPVPLFLAAEAVATVTSFLLSAFFPSANLAAACGLAYSLYLYPYVLCVAV 677
Db 708 NLLPSDPSVVFVLSVFAVVTILQCLISTLSFSAANLAAACGGIYFTLYPYVLCVAV 767
Qy 678 RDRLPAGGRVAASLLSPVAFGFCBSLALLBEEQGAQWNNVTGRT--ADVFSLAQVSGL 736
Db 768 QDYVGFTLKI FASLLSPVAFGCEYFALFEQIGVQWDLNLFESVEEDGNLTITSVM 827
Qy 737 LLLDAALYGLATWYLEAVCPGQYGIPEPWNPPFRSRYMGCPRPKSPACPPTPLD--PKYL 795
Db 828 MLFDTFLYGVMTWYIEAVFPQYGI PRPWYFPCTKSYWFEESDESKSHPSGNKRISEIC 887
Qy 796 VEEAPGLSPGVSVRSLEKRPGPSQPALRGLSLDPYOGHITAFIGHNCAAGTTLTSLIS 855
Db 888 MEEETHLKLGVSLQNLVYRDGMKAVDGLALNFBYEQIITSFLGHNGAGTTLTMSILT 947
Qy 856 GLFPPSGGSATILGHDRSSMAAIRPHLGVCPQYNVLFMDLTVDBHVMFYGRKLGLSAAV 915
Db 948 GLFPTSGTAYILGKD IRSEMTIRQNLGVCPQHNVLFMDLTVDBHIFYARLKLSEKH 1007
Qy 916 VGPEODRLLDVGL--VSKQSVQTRHLSGGMQRKLSVAIFVGGSQVWILDEPTAGVDPAS 974
Db 1008 VKAEMEQMALDVGLPSSKLSKTSQLSGGMQRKLSVALAFVGGSKVWILDEPTAGVDPYS 1067
Qy 975 RRGWELLKYREGTTLTSLTHLDEALLGDRVAVAGGRLLCCCGSLPLRRLHLSGVY 1034
Db 1068 RRGWELLKYRQGTTLTSLTHMDEADVLDGRTAIISHGKLCCCGVSSFLKNQGLTGY 1127
Qy 1035 LTLVKARLPLTN-----EKADTMEGSDVTROEKNKSGSQRVGTPTOLLALV 1082
Db 1128 LTLVKQDVESLSSCRNSSSTSVYLLKEDSVSQSSDAGLSDHSDTLTIDVSAISNLI 1187
Qy 1083 QHWVPGARVLELPHLVLVLPYTGADHGSFATLFRLEDTRIAELURLTGYGISDTSLEBI 1142
Db 1188 RKVSEARLVEDIGHELTYVLPYEAKEGAFVLEFHEIDRLSDLGISYISGSETTLEBI 1247
Qy 1143 FLKVEECAATDMEDSCGQHLCTGIAGLDVTLRLKMPPOETALE--NGEPGASAPETD 1200
Db 1248 FLKVAESGVDAETSDGTLPARNRRAF--DKQSCLRPFTEDDAADPNDSDIDPESRETD 1306
Qy 1201 QSGSPDAVG--RVQGWALTRQOLALLKRLFLAARRSRRLFAQIVLPALFVGLALVFSL 1258

```
Db 1307 LLSGMDKGSYQVKGWKLTOQFVALLWKLARRRKGFFAQIVLPAVFCVIALVPSL 1366
Qy 1259 IVPFPGHYPALRLSPMTYGAQVSPFSDAPGDPGRARLLALQEAQ-----LEE 1308
Db 1367 IVPFPGKPSLELQPMYNEQYTFVSDAPEDTGTLELLNALTDPGFGTRCMEGNIPD 1426
Qy 1309 PVOVSHSRPSAPVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAGPPP 1368
Db 1427 TPCQAGSEEWTTAPVQTIWDLFONGNWTWQNSPACQSSDKIKMLPVCPPCAGGLPP 1486
Qy 1369 POAVTSGSEVQNLGTGNLSDFLVKTYPRLVROGLTKKWNEVRYGGSIG-GRDPGLP 1427
Db 1487 PQRKQNTADILQDLTGRNISDYLVTYQIIAKSLNKIWNNEFRYGFSLGVSNTQALP 1546
Qy 1428 SGQELGRSVBELWALLSPGGLDRVLKNIITWAHSLDAQDSLKIWFNNKGWHSMAVF 1487
Db 1547 PSQEVNDAIKQMKHKLAKDSSADRLNSLGRFMTGLDTKNVKNVFNKKGWHAISFL 1606
Qy 1488 NRASNAILRALHPLPGPARHAHSITTLNHPNLNLTKEQLSEALMASSVDVLVSVCPFAMS 1547
Db 1607 NVINNAILRANLQGENPSHYGITAFNHPNLNLTQQLSEVALMTTSVDVLVSVCPFAMS 1666
Qy 1548 FVPASFVLVLEERVTRAKHLQMGLSPTLYMLGNFLWMCNVLVPACIVVLIFLAFQ 1607
Db 1667 FVPASFVFLIQERVSXAKHLQFTSGVKPVIYLSNFVDMCNVVPATVILIIIFCQQ 1726
Qy 1608 RAYVAPANLPALELLLLYGSITPLMPASFPSVPSTYAVVLTCINLFIGINGSMATF 1667
Db 1727 KSYVSSNLPVLALELLLLYGSITPLMPASFVKIPSTAYVVLTSVNLFIGINGSVATF 1786
Qy 1668 VLELFSQOKQEVSRILKQVLIPIPHFCLGRGLIDMYRNOAMADAFERLGRDQFQSPLRW 1727
Db 1787 VLELFTDKLNNINDILKSVELIPIPHFCLGRGLIDMYRNOAMADALRFENRFVPSLSW 1846
Qy 1728 EVYCKNLLAMVIOQLFLLTLLQHSQLLPOPRVSLPLGLCEDEVDAREERVVOGA 1787
Db 1847 DLVGRNLFAMAVEGVVFLITVLIIQYRFFIRPVNAKLSPLENDEDDVRERQRILDGG 1906
Qy 1788 TQGDVLVRLNLTKYVQRMPADVRLCLGIPPGCFGLLVNGAGKSTTFMTVGTDTLAS 1847
Db 1907 QONDILEIKELTKIYRKRKPAVDRI CVGIPPGCFGLLVNGAGKSTFKMLTGDITVT 1966
Qy 1848 RGEAVLAGHSVAREPSAAHLSMGVCPQSDAIFELLTGREHLELLARLGVPEAQVATAG 1907
Db 1967 RGDAFLKNSTLSIHEVHQNMGYCPQDAITELLTGREHVEFPALLRGVPEKEVGKVE 2026
Qy 1908 SGLARLGLSWYADRPAGTYSGGNKRKLATALVGDPAVVLDPBTPPTGMDPSARRFLWNS 1967
Db 2027 WAIRKGLVYGEKYGAGNYSGNKRKLSTAWALIGBPVFLDEPTTGMDPKARRFLWNC 2086
Qy 1968 LLAVVREGRSYMLTSHSMECEALCSRLAIWNGRFRCLGSPQLKGRFAAGHTLTLRVP 2027
Db 2087 ALSVVKSGRSVLTSHSMECEALCTRMAIWNGRFRCLGSPQLKGRFGDGYTIVRIA 2146
Qy 2028 AARS--OPAAAFVAEPFPGSELRAHGRRLRFPQLPPGRCALARVFGELAVHGAHVED 2085
Db 2147 GSNPDLKVPVDFGLAFPGSLVKEKRNMLQYQL-PSLSLSLARIFSILSOSKRLHIED 2205
Qy 2086 FSVSQTMLEEVLYFSKQCKDE--DTEQKEAGVGVDPAQGLQHPKRVSQFLDDPSTA 2142
Db 2206 YSVSQTLTDQVFNFAKDDQSDDDLKOLSLHKQTV-VDVAV-----LTFSLQDEKVK 2257
Qy 2143 ET 2144
Db 2258 ES 2259
```

RESULT 15

```
US-10-428-551-9
; Sequence 9, Application US/10428551
; Publication No. US20030229062A1
; GENERAL INFORMATION.
```

```
; APPLICANT: Ishida, Brian
; APPLICANT: Duncan, Keith
; APPLICANT: Bailey, Kathy
; APPLICANT: Kane, John
; APPLICANT: Schwartz, Daniel
; TITLE OF INVENTION: Treatments for Age Related-Macular Degeneration (AMD)
; FILE REFERENCE: P02351U53
; CURRENT APPLICATION NUMBER: US/10/428,551
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/415,864
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US 60/340,498
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 2261
; TYPE: PRT
; ORGANISM: Human
; US-10-428-551-9

Query Match 51.8%; Score 5768.5; DB 14; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;

Qy 1 MATWTQMLLLWKNFMYRRRQPVQLLVLLWPLPLFFILVAVRHSHPHHECHFFNKP 60
Db 1 MACWPQLRLLWKNLFRRRQTCQLLLENAVPLFIILLISVRLSYPPYEQHECHFPNKA 60
Qy 61 LPSAGTVPMLOGLICNVNNTCPQLTPGEBPGRLSNFNDLSVRLSLADARTVLGGASAH 120
Db 61 MPSAGTLPWVGIIICNANNPCFRYPITPGAPVGVGNFNKSIVARLFSADARLLYSQDT 120
Qy 121 TLAGLGLIATLAAARSTAQ-----PQTKOSPLEPML-- 154
Db 121 SMKDMRKLRTLQIKKSSSLKQLDFVDNETFSGFLYHNLSPKSTVDMKLADVLH 180
Qy 155 -----DVAEL-----LTSLL 164
Db 181 KVFLOQYQLHLTSLNGSKSEMIQLDQEVSLCGLPREKLAARVLRNSMDILKPI 240
Qy 165 RT--ESLGLAQAQAPLHSLLEAAEDLAQELLALRSLVELR--ALLQRPRTSGGLE 219
Db 241 RTLNSTSPFPSKELAEATKLLHSLGTLAQELSMRSWSMDRQVMPFLTNVNSSSTQI 300
Qy 220 ---LSEALCSVRGPSSTVGPSLWYBASDLMELVG---QEPESALPDSSSPACSLIG 272
Db 301 YQAVSRIVCGHPEGGLIKILNWDNNYKALFGNGTEEDAETFYDNSTTPYCNLDMK 360
Qy 273 ALDSHPLRLLWRRLKPLILGLKLLPAPDTPFKLMAQVNRTEELTLDRDREVWELG 332
Db 361 NLESSPLSRIIWKALPPLLVGKILYTPDTPATRVMAEVNKTFOELAVFHDLEGWBEELS 420
Qy 333 PRITFMNDSSNAMLQRLLOMDEGR-ROPREGRD-----HMEALRSFLDP 379
Db 421 PKIWTMNSQEMDLVRMLLDSDRDNHFWEQQLDGLDWTADQIVAFIAKPEDVQS---- 476
Qy 380 GSGG--YSWODAHADVHLVGTILGRVTECLSDKLEAAPSEAAALVSALQLLAHFRWAG 437
Db 477 -SNGSVYTWEAFNETNQARTISRFMECVNLKLEPIATEVWLINKSMELLDERKFWAG 535
Qy 438 VVFLGEDSSDPTEHPTDPLGPGHVRIKIRMDIDVTRTNKIRDRFWDGCPADPLDLR 497
Db 536 IVFTGTPGSIELPH-----HVYKIRMDIDNVERTNKIKDGYWDPGPRADPFEDMR 587
Qy 498 YVMGGFVYLQDLVERAAVRVLSGANPRAGLYLOQMPYPCVDDVFLRVLRSPLFLTLA 557
Db 588 YVMGGFAYLQDVVEQAIIRVLTGTEKTYVMQMPYPCVDDIFLVRNSRNPFLWTLA 647
Qy 558 WIYSVTLTVKAVVREKTRLRDTRMRAMGLSRVILWGLFSLGCPFLLSAALLVLVLKLG 617
Db 648 WIYSVAVIIRKGIIVEKEARKETWRIMGLDNLWFSWFISSLTPLIVSAGLLVLVLKLG 707
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 10:47:08 ; Search time 180.939 Seconds
(without alignments)
4254.650 Million cell updates/sec

Title: US-09-995-542-5
Perfect score: 11143
Sequence: 1 MAFWTQLMLLWKNFMYRR.....QHPKRVQFLDDPSTAETVL 2146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 359729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	11143	100.0	2146	6	Abg72696 Human ATP
2	11130	99.9	2146	4	Aau04483 Human PD-
3	11130	99.9	2180	5	Aaol4210 Human tra
4	11128	99.9	2146	5	Abp52096 Homo sapi
5	11128	99.9	2146	6	Abu08466 Human ABC
6	11107	99.7	2144	5	Aau09174 Human tra
7	11107	99.7	2144	7	Add37429 Human tra
8	10639.5	95.5	2059	6	Abu54629 Human NOV
9	10264	92.1	2008	6	Abu08464 Amino aci
10	10171.5	91.3	1993	6	Abu08465 Amino aci
11	9460.5	84.9	1873	4	Aau04484 Human PD-
12	8613.5	77.3	2167	6	Abg72695 Mouse ATP
13	7596	68.2	1550	6	Abg72697 Human ATP
14	5769.5	51.8	2261	3	Aab38111 Human ABC
15	5769.5	51.8	2261	3	Aab38114 Human ABC
16	5768.5	51.8	2261	3	Aab38117 Human ABC
17	5768.5	51.8	2261	3	Aab38115 Human ABC
18	5768.5	51.8	2261	3	Aab38109 Human ABC
19	5768.5	51.8	2261	3	Aab38082 Human ABC
20	5768.5	51.8	2261	3	Aab38112 Human ABC
21	5768.5	51.8	2261	4	Aab71749 Human ABC
22	5768.5	51.8	2261	4	Aab31361 Amino aci
23	5768.5	51.8	2261	4	Aab31365 Amino aci
24	5768.5	51.8	2261	6	Abu11899 Human ATP
25	5768.5	51.8	2261	6	Abp62034 Human ABC

26	5768.5	51.8	2261	6	ABR62033 Human ABC
27	5768.5	51.8	2261	7	ADP65173 Human ATP
28	5767.5	51.8	2261	3	AAB38116 Human ABC
29	5766.5	51.7	2261	3	AAB38113 Human ABC
30	5764.5	51.7	2261	3	AAB38110 Human ABC
31	5764.5	51.7	2261	3	AAB38105 Human ABC
32	5764.5	51.7	2261	4	AAB31362 Amino aci
33	5764.5	51.7	2261	4	AAB31366 Amino aci
34	5760.5	51.7	2261	4	AAB31363 Amino aci
35	5760.5	51.7	2261	4	AAB31367 Amino aci
36	5758	51.7	2260	3	AAB38106 Human ABC
37	5756.5	51.7	2261	3	AAB38104 Human ABC
38	5752.5	51.6	2261	5	ABR83120 Polymorph
39	5752.5	51.6	2261	5	ABR83116 Polymorph
40	5752.5	51.6	2261	5	ABR83124 Polymorph
41	5752.5	51.6	2261	5	ABR83122 Polymorph
42	5751.5	51.6	2259	3	AAB38107 Human ABC
43	5751.5	51.6	2261	4	AAM50228 Human ATP
44	5751.5	51.6	2261	4	AAE13022 Human ATP
45	5751.5	51.6	2261	4	AAU02176 Human ABC

ALIGNMENTS

RESULT 1
ABG72696
ID ABG72696 standard, protein; 2146 AA.
XX AC ABG72696;
XX AC
XX 10-MAR-2003 (first entry)
XX DE Human ATP-binding cassette transporter-like protein, ABCL.
XX KW Human; ATP-binding cassette transporter-like protein; ABCL;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropenia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..46
Protein	/label= Signal_peptide 47..2146 /label= Mature_ABCL
US2002127647-A1.	
12-SEP-2002.	
28-NOV-2001; 2001US-00995542.	
28-NOV-2000; 2000US-0253520P.	

(SHUT/) SHUTTER J.
(ULIA/) ULIAS L.
Shutter J, Ulia L;
WPI; 2003-147394/14.
N-FSDS; ABX14666.
Novel ATP-binding cassette transporter-like polypeptides and polynucleotides useful for diagnosing, preventing, treating disorders

PT involving immune, nervous system, thyroid, hypothalamus and impaired
FT transport of lipids.

XX Claim 13; Fig 2; 149pp; English.

XX The invention relates to an isolated murine and human ATP-binding
CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
CC encoded by the DNA insert in AtCC Deposit Nos PTA-3109, PTA-3110 or PTA-
CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
CC vectors, host cells, ABCL binding agents, a selective binding agent or
CC its fragment comprising at least one complementarity determining region
CC (CDR) with specificity for ABCL which (produced by immunising an animal
CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
CC fusion polypeptide, a device comprising a membrane suitable for
CC implantation (permeable to the protein and impermeable to materials
CC detrimental to the cells, and cells encapsulated within the membrane)
CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
CC and modulators are useful for the diagnosis and/or treatment of diseases
CC and conditions involving impaired transport of lipids, including
CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
CC involving functional and trophic disturbances of the nervous system such
CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
CC fibrosis, conditions involving multidrug resistance, conditions involving
CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
CC neutropenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies; autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL

XX Sequence 2146 AA;

Query Match 100.0%; Score 11143; DB 6; Length 2146;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFWTQMLLWKMFYRRRRQPVQLLVLLWPLFLFVLVAVRSHHPPLBHHCHFFNKP 60
DB 1 MAFWTQMLLWKMFYRRRRQPVQLLVLLWPLFLFVLVAVRSHHPPLBHHCHFFNKP 60
QY 61 LPSAGTVPWLOGLCNVNNTCFPOLTCGEEPRGRLSNFNDSLVSRLLADARTVLGGASAH 120
DB 61 LPSAGTVPWLOGLCNVNNTCFPOLTCGEEPRGRLSNFNDSLVSRLLADARTVLGGASAH 120
QY 121 TLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTSLGALGQAEPL 180
DB 121 TLAGLGLIATLRAARSTAOPTKQSPLEPPMLDVAELTSLRTSLGALGQAEPL 180
QY 181 HSLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELSEALCSVRGSPSTVGPSLNW 240
DB 181 HSLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELSEALCSVRGSPSTVGPSLNW 240
QY 241 YEASDLMELVQEPESALPDSSLSFACSELTGALDHPHLSRLWRLLKPLTLGLKLLFAPD 300
DB 241 YEASDLMELVQEPESALPDSSLSFACSELTGALDHPHLSRLWRLLKPLTLGLKLLFAPD 300
QY 301 TPFTFKLMAQVNRTEELTLRDVREVMELGPRIFTFMNDSNVNMLQRLLOQWDEGRR 360
DB 301 TPFTFKLMAQVNRTEELTLRDVREVMELGPRIFTFMNDSNVNMLQRLLOQWDEGRR 360
QY 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLGTVTLGRVTECLSDKLEAAPSEAA 420
DB 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLGTVTLGRVTECLSDKLEAAPSEAA 420
QY 421 LVSRALQLLABHRFWAGVVFGLPDSDDPTHEPTDLPGLPGHVRKIRMDIDVTRTNKIR 480
DB 421 LVSRALQLLABHRFWAGVVFGLPDSDDPTHEPTDLPGLPGHVRKIRMDIDVTRTNKIR 480
QY 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPICYDD 540

DB 481 DRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYPICYDD 540
QY 541 VFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLRDMRAMGLSRVAVLWLGWFLSCL 600
DB 541 VFLRVLSRSLPLFLTLAWIYSVTLTKAVVREKETRLRDMRAMGLSRVAVLWLGWFLSCL 600
QY 601 GPFLLSALLVIVLKLGDILPYSHPGVWFLFLAAFAVATVTSQFLLSAFSSRANLAAACG 660
DB 601 GPFLLSALLVIVLKLGDILPYSHPGVWFLFLAAFAVATVTSQFLLSAFSSRANLAAACG 660
QY 661 GLAYFSLYLPYLVCVARDRLPAGGRVAAASLLSPVAFGFCESLALLEEGEGEAQHNVG 720
DB 661 GLAYFSLYLPYLVCVARDRLPAGGRVAAASLLSPVAFGFCESLALLEEGEGEAQHNVG 720
QY 721 TRPTADVLSLAQVSGLLLDAAALYATWYLEAVCPQYGIPEPWNPPFRSSVWCGRPPP 780
DB 721 TRPTADVLSLAQVSGLLLDAAALYATWYLEAVCPQYGIPEPWNPPFRSSVWCGRPPP 780
QY 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLDFYOGHITAF 840
DB 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQALRGSLDFYOGHITAF 840
QY 841 GHNGAGKTTTSLISGLFPFPGSAGFTLGHDRSSMAAIRPHLGVCPQYVNLFDMLTVDE 900
DB 841 GHNGAGKTTTSLISGLFPFPGSAGFTLGHDRSSMAAIRPHLGVCPQYVNLFDMLTVDE 900
QY 901 HVMFYGRKLKLSAAVVGPEQDRLLQDVGLSVKQSVQTRHLSGGWQKLSVAIAFVGGSQV 960
DB 901 HVMFYGRKLKLSAAVVGPEQDRLLQDVGLSVKQSVQTRHLSGGWQKLSVAIAFVGGSQV 960
QY 961 VIIDEPAGVDPASRRGIWELLILKYREGRTLILSTHHLDEALLGDRVAVVAGRLCCCG 1020
DB 961 VIIDEPAGVDPASRRGIWELLILKYREGRTLILSTHHLDEALLGDRVAVVAGRLCCCG 1020
QY 1021 SPLFLRRHLGSGYVLTLVKARLPLTNEKADTMEGSDVTRQEKNGSQSGSRVGTTPOLLA 1080
DB 1021 SPLFLRRHLGSGYVLTLVKARLPLTNEKADTMEGSDVTRQEKNGSQSGSRVGTTPOLLA 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAEURLTGYGSDTSLE 1140
DB 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRLAEURLTGYGSDTSLE 1140
QY 1141 EIFLKVVEECAADTMEGSCGQHLCTGIAGLDTLRLKMPPOETALENGEPAGSAPETD 1200
DB 1141 EIFLKVVEECAADTMEGSCGQHLCTGIAGLDTLRLKMPPOETALENGEPAGSAPETD 1200
QY 1201 QSGSPDAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFLPAQIVLPALFVGLALVFSLI 1260
DB 1201 QSGSPDAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFLPAQIVLPALFVGLALVFSLI 1260
QY 1261 PPFQHYPALRLSPMTYGAQVSVFSEADAPDGPGRARLLLEALLQAGLEPEPQVSHSRFSA 1320
DB 1261 PPFQHYPALRLSPMTYGAQVSVFSEADAPDGPGRARLLLEALLQAGLEPEPQVSHSRFSA 1320
QY 1321 PEYPAEYAKVLASGNWTPESPSPACQCSQPGARLLPDCPAAAGPPPPQAVTSGSGVWQ 1380
DB 1321 PEYPAEYAKVLASGNWTPESPSPACQCSQPGARLLPDCPAAAGPPPPQAVTSGSGVWQ 1380
QY 1381 NLTGRLNLSDFLVKTYPRLVQGLTKKWNVEYVGGFSLGGRDPGLSPGQELGRSVEELW 1440
DB 1381 NLTGRLNLSDFLVKTYPRLVQGLTKKWNVEYVGGFSLGGRDPGLSPGQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNLTAWAHSLLDAQSDLSKTIWNNKNGHSMVAFVNRASNAI 1500
DB 1441 ALLSPLPGGALDRVLKNLTAWAHSLLDAQSDLSKTIWNNKNGHSMVAFVNRASNAI 1500
QY 1501 PGPARHAHSITTLNHPNLTKQOLSEALMASVDVLSICVVFAMSVFASFTVLVLEE 1560
DB 1501 PGPARHAHSITTLNHPNLTKQOLSEALMASVDVLSICVVFAMSVFASFTVLVLEE 1560
QY 1561 RVTRAKHLQMLGGLSPTLYWLGNFMDMCMNLYLPACIVLILFLAQOQAYVAPANL 1620

Db 1561 RVTRAKHLQMGSLPTLYWLNFLNLCMNYLPVAPICVVLIFLAFQRAYVAPANLPALL 1620
Qy 1621 LLLLYGWSITPLMYPASFFRSVPSTAYVVLTCINLFTGINGSMATFVLELPSDOKLOEV 1680
Db 1621 LLLLYGWSITPLMYPASFFRSVPSTAYVVLTCINLFTGINGSMATFVLELPSDOKLOEV 1680
Qy 1681 SRLKQVFLIPPHFCGLRGLDMVRNQAMADAFERLGRQFQSPLRWEVVGKLLAMVIQ 1740
Db 1681 SRLKQVFLIPPHFCGLRGLDMVRNQAMADAFERLGRQFQSPLRWEVVGKLLAMVIQ 1740
Qy 1741 GFLFLFTLLQHRSQLLPQPRVRSPLIGEDEDVARERERVVGATQGDVLRNLTK 1800
Db 1741 GFLFLFTLLQHRSQLLPQPRVRSPLIGEDEDVARERERVVGATQGDVLRNLTK 1800
Qy 1801 VYRGORMPAVDRLCIGIPGECFGLLVGNVAGKTTSTPRWVTGDTLASGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRLCIGIPGECFGLLVGNVAGKTTSTPRWVTGDTLASGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1920
Db 1861 EPSAAHLSMGYCQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSGLARLGLSWYAD 1920
Qy 1921 RPAGTYSNGNKKLATALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGRSVWL 1980
Db 1921 RPAGTYSNGNKKLATALVGDPAVFLDEPTTGMDPSARRFLWNSLLAVVREGRSVWL 1980
Qy 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Db 1981 TSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAFVAA 2040
Qy 2041 EFPGSSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEBVFLYF 2100
Db 2041 EFPGSSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEBVFLYF 2100
Qy 2101 SKDQKDEBTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146
Db 2101 SKDQKDEBTEQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAEVTL 2146

RESULT 2

AAU04483
ID AAU04483 standard; protein; 2146 AA.

AC AAU04483;

XX 26-SEP-2001 (first entry)

DT Human PD-ATP-binding cassette (PD-ABC) protein form #1.

XX PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; Tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.

XX Homo sapiens.

OS WO200153490-A1.

XX 26-JUL-2001.

XX 23-JAN-2001; 2001WO-US002191.

XX 24-JAN-2000; 2000US-0177889P.

XX 30-JUN-2000; 2000US-0215405P.

XX (WARN) WARNER LAMBERT CO.

XX Johns MA, Tafuri SR, Wang M;

XX WPI; 2001-442259/47.

DR N-PSDB; AAS08706.

XX New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
PT of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.

PS Claim 10; Page 54-62; 77pp; English.

XX The sequence represents human PD-ATP-binding cassette (PD-ABC) protein
form 1. PD-ABC maps to chromosome 19p13.3 and is expressed in various
tissues including spleen, thymus, peripheral blood leukocytes, bone
marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to
diagnose and treat cardiovascular disorders, inflammatory disorders,
dyslipidaemia, epilepsy, diseases related to abnormal calcium flux,
coronary artery disease, Tangier's disease, familial high-density
lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
insulin resistance, obesity, alcoholism, retinal degeneration,
hypertension and vascular disease. The sequences are also used in drug
screening assays

SQ Sequence 2146 AA;

Query Match 99.9%; Score 11130; DB 4; Length 2146;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2144; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAFTQMLLLKMKFMYRRROPVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60

Db 1 MAFTQMLLLKMKFMYRRROPVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60

Qy 61 LPSAGTVPMWQLGICNVNNTCFPQITPGEEFGRSLNFNDSLVSRLLADARTVLGASAH 120

Db 61 LPSAGTVPMWQLGICNVNNTCFPQITPGEEFGRSLNFNDSLVSRLLADARTVLGASAH 120

Qy 121 TLAGLGLIATLRAARSTAQPTKQSPLEPPMLDVAELLTSLRTESIGLALGOAQBPL 180

Db 121 TLAGLGLIATLRAARSTAQPTKQSPLEPPMLDVAELLTSLRTESIGLALGOAQBPL 180

Qy 181 HSLEAEADLAQELLALRSIVELRALLQRPSTGPLELLSEALCSVRGSPSTVGPSSLNW 240

Db 181 HSLEAEADLAQELLALRSIVELRALLQRPSTGPLELLSEALCSVRGSPSTVGPSSLNW 240

Qy 241 YEASDLMELVGQEPESALPDSSLPACSELIGALDHPSLRLLMRLLKPLILGLLFPD 300

Db 241 YEASDLMELVGQEPESALPDSSLPACSELIGALDHPSLRLLMRLLKPLILGLLFPD 300

Qy 301 TPFTKRLMAQVNRTFEELTLRLDVRVWEMLGPRITFTFMDSSNVAMLQRLLOMDEGR 360

Db 301 TPFTKRLMAQVNRTFEELTLRLDVRVWEMLGPRITFTFMDSSNVAMLQRLLOMDEGR 360

Qy 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTLCGRVTECLSLDKLEAPSEAA 420

Db 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGLVGTLCGRVTECLSLDKLEAPSEAA 420

Qy 421 LVSRALQLLAEHRFWAGVVFGLPESDSDPTSEHTPDLPDGPVHRKIRMDIDVVTTRNKR 480

Db 421 LVSRALQLLAEHRFWAGVVFGLPESDSDPTSEHTPDLPDGPVHRKIRMDIDVVTTRNKR 480

Qy 481 DRFWDPGPADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYCYDD 540

Db 481 DRFWDPGPADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAGLYLOMPYCYDD 540

Qy 541 VFLRVLSRSLPLFTLTAWIYSVTLTVKAVVREKETRLRDTWRAMGLSAVWLWGLFSLCL 600

Db 541 VFLRVLSRSLPLFTLTAWIYSVTLTVKAVVREKETRLRDTWRAMGLSAVWLWGLFSLCL 600

Qy 601 GPFLLSAALLVLVLKGLDILFYSHPGVVFLEAFAVATVTSFLLSAFFSRANLAAACG 660

Db 601 GPFLLSAALLVLVLKGLDILFYSHPGVVFLEAFAVATVTSFLLSAFFSRANLAAACG 660

Qy 661 GLAYFSYLYPYVLCVAVWRDLRPLAGRVAASLLSPVAFGFCESIALLEEQEGQAHNVG 720

Db 661 GLAYFSYLYPYVLCVAVWRDLRPLAGRVAASLLSPVAFGFCESIALLEEQEGQAHNVG 720

QY 721 TRPTADVSLAQVSGLLLDALYGLATWYLEAVCPQYGIPEPWNPPFRSSWCGPRPP 780
DB 721 TRPTADVSLAQVSGLLLDALYGLATWYLEAVCPQYGIPEPWNPPFRSSWCGPRPP 780
QY 781 KSPAPCPTPLDPKVLVEAPPGSLSPGVSRSLEKRFPSPQALRGSLDFYQGHITAF 840
DB 781 KSPAPCPTPLDPKVLVEAPPGSLSPGVSRSLEKRFPSPQALRGSLDFYQGHITAF 840
QY 841 CHNGAGKTTTILSILSGLFPSPGSAFILGHDVRSMAAIRPHLGVCQYNNVLFMDLTVDE 900
DB 841 CHNGAGKTTTILSILSGLFPSPGSAFILGHDVRSMAAIRPHLGVCQYNNVLFMDLTVDE 900
QY 901 HVWFYGRKGLSAAVGPEDRLLODVLGYSKQSVQTRHLSGGWOKLSVAIAFVGGSOV 960
DB 901 HVWFYGRKGLSAAVGPEDRLLODVLGYSKQSVQTRHLSGGWOKLSVAIAFVGGSOV 960
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1020
DB 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDDEALLGDRVAVVAGRLCCCG 1020
QY 1021 SPLFRRLHSGYYLTLVKARLPLTTNEKADTMEGSVDTQREKNGSQSRVGTPOLLA 1080
DB 1021 SPLFRRLHSGYYLTLVKARLPLTTNEKADTMEGSVDTQREKNGSQSRVGTPOLLA 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAELRLTGYSIDTSLE 1140
DB 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLRELDTRLAELRLTGYSIDTSLE 1140
QY 1141 EIFLKVVVEECADTMEDGSCGQHLCTGIAGLDVTLRLKMPPOQTALENGEPAGSAPETD 1200
DB 1141 EIFLKVVVEECADTMEDGSCGQHLCTGIAGLDVTLRLKMPPOQTALENGEPAGSAPETD 1200
QY 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARRSRRLGFAQIYVLPALFVGLALVFSLIV 1260
DB 1201 QGSGPDVAVRQGWALTRQQLQALLKRLFLARRSRRLGFAQIYVLPALFVGLALVFSLIV 1260
QY 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGPGARLLEALLQEALEBPVQHSSHRFSA 1320
DB 1261 PPFQHYPALRLSPMYGAQVSFFSEDAPGPGARLLEALLQEALEBPVQHSSHRFSA 1320
QY 1321 PEVPAEVAKVLASGNWTPESPSPACQSQCARLRLPDCPAAAGPPPPQAVTGSGEVWQ 1380
DB 1321 PEVPAEVAKVLASGNWTPESPSPACQSQCARLRLPDCPAAAGPPPPQAVTGSGEVWQ 1380
QY 1381 NLTGRLNSDFLVKTYPRLVQGLKTKKWNVEVRYGFSLGGRDPLSPGQELGRSVEELW 1440
DB 1381 NLTGRLNSDFLVKTYPRLVQGLKTKKWNVEVRYGFSLGGRDPLSPGQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTAWAHSLEDAQDSLKIFWNNKGHSMVAFVYNRASNAILLRAHLP 1500
DB 1441 ALLSPLPGGALDRVLKNTAWAHSLEDAQDSLKIFWNNKGHSMVAFVYNRASNAILLRAHLP 1500
QY 1501 PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSCVWFAMSFVPASFTVLVLEE 1560
DB 1501 PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLVSCVWFAMSFVPASFTVLVLEE 1560
QY 1561 RVTRAKHLQMGGLSPTLYWLNGLFDMCNLYVPACIVLIFLAFQORAYVAPANLPAH 1620
DB 1561 RVTRAKHLQMGGLSPTLYWLNGLFDMCNLYVPACIVLIFLAFQORAYVAPANLPAH 1620
QY 1621 LLLLYGWSITPLMYPASFPSPVSTAYVVLVTCINLFIGINGSMATFVLELFSQKLOEV 1680
DB 1621 LLLLYGWSITPLMYPASFPSPVSTAYVVLVTCINLFIGINGSMATFVLELFSQKLOEV 1680
QY 1681 SRIILKQVFLIIPHFCLGRLIDMVNRQAMADAPERLGDROFQSPLRWEVVGKLLAMVIQ 1740
DB 1681 SRIILKQVFLIIPHFCLGRLIDMVNRQAMADAPERLGDROFQSPLRWEVVGKLLAMVIQ 1740
QY 1741 GFLFLLFTLLQHRSQLLPQPRVRSPLLGGEEDVARERVRVQGTQGDVLVLRNLTK 1800
DB 1741 GFLFLLFTLLQHRSQLLPQPRVRSPLLGGEEDVARERVRVQGTQGDVLVLRNLTK 1800
QY 1801 VYRQORPAMDRLCLGIPPGCEFGLLGVNGAGKTSIFRMVTGDTLASRGEAVLAGHSVAR 1860

DB 1801 VYRQORPAMDRLCLGIPPGCEFGLLGVNGAGKTSIFRMVTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLSMGYCQSDAIPELLTGRHELELLARLGRVPEAQVAQTAGSLARLGLSWYAD 1920
DB 1861 EPSAAHLSMGYCQSDAIPELLTGRHELELLARLGRVPEAQVAQTAGSLARLGLSWYAD 1920
QY 1921 RPACTYSGGNKRLATALLVGDPAVVFLDEPTTGMDPSARRELWNSLLAVVREGRSVML 1980
DB 1921 RPACTYSGGNKRLATALLVGDPAVVFLDEPTTGMDPSARRELWNSLLAVVREGRSVML 1980
QY 1981 TSHSMECEALCSRLAIMVNGRFRCLSGPOHLKGRFAAGHTLTLRVPAPARSQPAARFVAA 2040
DB 1981 TSHSMECEALCSRLAIMVNGRFRCLSGPOHLKGRFAAGHTLTLRVPAPARSQPAARFVAA 2040
QY 2041 EPPGSELREAHGRLRFQPPGRCALARVFGELAVHGAHGVDFSVSQTMLVEEVFLYF 2100
DB 2041 EPPGSELREAHGRLRFQPPGRCALARVFGELAVHGAHGVDFSVSQTMLVEEVFLYF 2100
QY 2101 SKDQKDEDETEEQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL 2146
DB 2101 SKDQKDEDETEEQKEAGVGVDPAPGLQHPKRVSOQLDDPSTAETVL 2146
RESULT 3
AAO14210
ID AAO14210 standard; protein; 2180 AA.
XX AAO14210;
AC AAO14210;
DT 03-MAY-2002 (first entry)
XX Human transporter and ion channel TRICH-27.
DE Human; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder;
KW cell proliferative disorder; neuroprotective; nootropic;
KW cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
KW gene therapy.
XX Homo sapiens.
OS Homo sapiens.
XX WO200204520-A2.
PN 17-JAN-2002.
XX 05-JUL-2001; 2001WO-US021448.
XX 07-JUL-2000; 2000US-0216547P.
PR 14-JUL-2000; 2000US-0218232P.
PR 21-JUL-2000; 2000US-0220112P.
PR 28-JUL-2000; 2000US-0221839P.
XX (INCY-) INCYTE GENOMICS INC.
XX Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J;
PI Patterson C, Gandhi AR, Hafalia AJA, Tribouley CM, Wailia NK;
PI Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
PI Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan FA;
PI Kearney L, Thangavelu K, Das D, Policky JL;
XX WPI; 2002-205969/26.
DR N-PSDB; AAL44693.
XX New human transporters and ion channel polypeptides for diagnosing,
PT treating or preventing transport, neurological, muscle, immunological and
PT cell proliferative disorders.
XX Claim 1; Page 188-192; 230pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of human transporter and ion channel proteins, designated TRICH-1-

CC TRICH-32. The sequences can be used in the treatment of transport, CC neurological, muscle, immunological and cell proliferative disorders. The CC present sequence is a protein of the invention

XX Sequence 2180 AA;

Query Match 99.9%; Score 11130; DB 5; Length 2180;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2143; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY	1	MAFWTQLMLLKWNFMYRRQPVQLLVLLWPLFLFFILVAVRHSHPPLEHHECHFPNKP	60
DB	35	MAFWTQLMLLKWNFMYRRQPVQLLVLLWPLFLFFILVAVRHSHPPLEHHECHFPNKP	94
QY	61	LPSAGTVPMLOGLICNVNNTCPQPTPGEPGRKLSNFNDLSVRLIADARTVLGGASAH	120
DB	95	LPSAGTVPMLOGLICNVNNTCPQPTPGEPGRKLSNFNDLSVRLIADARTVLGGASAH	154
QY	121	TLAGLGKLIATLRAARSTAQOPTKQSPLEPPMLDVAELLTSLRTESLGLAQOEBPL	180
DB	155	TLAGLGKLIATLRAARSTAQOPTKQSPLEPPMLDVAELLTSLRTESLGLAQOEBPL	214
QY	181	HSLLAEADLAQELLALRSLSVLRALLQRPRTSGTSGPLELLSEALCSVRGPSTVGP	240
DB	215	HSLLAEADLAQELLALRSLSVLRALLQRPRTSGTSGPLELLSEALCSVRGPSTVGP	274
QY	241	YEASDLMELVCGEPESALPDSSLPACSELIGALDHPRLRLRLKPLILGKLLFAPD	300
DB	275	YEASDLMELVCGEPESALPDSSLPACSELIGALDHPRLRLRLKPLILGKLLFAPD	334
QY	301	TPFTKRLMAQVNRTPFEELTLRDREVWEMLGPRIFTFMNDSSNVAMLQRLLOMDEGR	360
DB	335	TPFTKRLMAQVNRTPFEELTLRDREVWEMLGPRIFTFMNDSSNVAMLQRLLOMDEGR	394
QY	361	QRPGRDHMEALRSFLDPGSGCYQWODAHVGLVGTLRVTECLSDLKLEAAPSEAA	420
DB	395	QRPGRDHMEALRSFLDPGSGCYQWODAHVGLVGTLRVTECLSDLKLEAAPSEAA	454
QY	421	LVSRALQLLAEHRFWAGVVFVLPEDSSDPTHTPTDLPDGHVRIKIRMDIDVVTNKIR	480
DB	455	LVSRALQLLAEHRFWAGVVFVLPEDSSDPTHTPTDLPDGHVRIKIRMDIDVVTNKIR	514
QY	481	DRFWDGPAADPLTLRLYVWGVFVYLQDLVERAAVRLSGANPRAGLYLQMPYCYDD	540
DB	515	DRFWDGPAADPLTLRLYVWGVFVYLQDLVERAAVRLSGANPRAGLYLQMPYCYDD	574
QY	541	VPLRVLSRLPLFLTLAMITYSVTLTKAVVREKETRLRDTMRAMGLRAVLWGLWFLSCL	600
DB	575	VPLRVLSRLPLFLTLAMITYSVTLTKAVVREKETRLRDTMRAMGLRAVLWGLWFLSCL	634
QY	601	GPFLLSAALLVLVKGDLIPYSHGQVWFLFLAAFAVATVTSFLLSAFFSRANLAAACG	660
DB	635	GPFLLSAALLVLVKGDLIPYSHGQVWFLFLAAFAVATVTSFLLSAFFSRANLAAACG	694
QY	661	GLAYFSLYPLVYLCAVMDRLPAGGRVAASLLSPVAFGFCESLALLEQEGEAQWNVG	720
DB	695	GLAYFSLYPLVYLCAVMDRLPAGGRVAASLLSPVAFGFCESLALLEQEGEAQWNVG	754
QY	721	TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYIGIPEPNWFFRYSYWCGRPP	780
DB	755	TRPTADVSLAQVSGLLLLDAALYGLATWYLEAVCPQYIGIPEPNWFFRYSYWCGRPP	814
QY	781	KSPAPCPTPLDPKVLVEAPPGSLPGSVRSLEKXFPQSPQALRGLSLDFYQGHITAF	840
DB	815	KSPAPCPTPLDPKVLVEAPPGSLPGSVRSLEKXFPQSPQALRGLSLDFYQGHITAF	874
QY	841	GHNGAGKTTTILSGLPSPSGSAF ILCHDVSRSSMAAIRPHLGYCPQYVNLFDMLTVDE	900
DB	875	GHNGAGKTTTILSGLPSPSGSAF ILCHDVSRSSMAAIRPHLGYCPQYVNLFDMLTVDE	934
QY	901	HWYFYGRLLKLSAAVVGPEQDRLLQDVLGVLSKQSVQTRHLSCGMQRKLSVAIAFYGGSQV	960
DB	935	HWYFYGRLLKLSAAVVGPEQDRLLQDVLGVLSKQSVQTRHLSCGMQRKLSVAIAFYGGSQV	994

QY	961	VILDEPTAGVDPASRRGIWELLKYREGRTLLSTLTHLDEALLGDRVAVVAGRLCCCG	1020
DB	995	VILDEPTAGVDPASRRGIWELLKYREGRTLLSTLTHLDEALLGDRVAVVAGRLCCCG	1054
QY	1021	SPLFRRHLGSGYYITLVKARLPLTTNEKADTDMEGSVDTROEKXGSGSVGTVPQLLA	1080
DB	1055	SPLFRRHLGSGYYITLVKARLPLTTNEKADTDMEGSVDTROEKXGSGSVGTVPQLLA	1114
QY	1081	LVQHWVPGARLVPELPHLVLVLPYTGADGSPATLPRELDTLAEALRTGIGSDTSLE	1140
DB	1115	LVQHWVPGARLVPELPHLVLVLPYTGADGSPATLPRELDTLAEALRTGIGSDTSLE	1174
QY	1141	EFLFVVECAADTDMEDGSCQHLCTGTAGLDVTLRLKMPQOETALENGEPAGAPETD	1200
DB	1175	EFLFVVECAADTDMEDGSCQHLCTGTAGLDVTLRLKMPQOETALENGEPAGAPETD	1234
QY	1201	QSGSDPAVGRVQGWALTQQLQALLKRLFLARRSRRLFAQIVLPALFVGLALVFLTLV	1260
DB	1235	QSGSDPAVGRVQGWALTQQLQALLKRLFLARRSRRLFAQIVLPALFVGLALVFLTLV	1294
QY	1261	PPFGHYPALRLSPTMYGAQVFFSDEADPGDGRARLLLEALQEGLEIEPPVQHSRPSA	1320
DB	1295	PPFGHYPALRLSPTMYGAQVFFSDEADPGDGRARLLLEALQEGLEIEPPVQHSRPSA	1354
QY	1321	PEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVEVQ	1380
DB	1355	PEVPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVEVQ	1414
QY	1381	NLTGNLSDFLVKTYPRLVRQGLTKKWNVNRYGFSLGGRDPLGSGQELGRSVEELW	1440
DB	1415	NLTGNLSDFLVKTYPRLVRQGLTKKWNVNRYGFSLGGRDPLGSGQELGRSVEELW	1474
QY	1441	ALLSPLPGALDRVLKNTAWAHSDDAODSLKIFNFKGWSHWAFAVRASNAIIRAHLP	1500
DB	1475	ALLSPLPGALDRVLKNTAWAHSDDAODSLKIFNFKGWSHWAFAVRASNAIIRAHLP	1534
QY	1501	PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLSICVVFAMSVFVPASFTLVLEE	1560
DB	1535	PGPARHAHSITTLNHPNLNLTKEQSEALMASSVDVLSICVVFAMSVFVPASFTLVLEE	1594
QY	1561	RVTRAKHLQMLGSLPTLYLWLNFCNLYLWPCIVVLIFLAFQQRAYVAPANLPALL	1620
DB	1595	RVTRAKHLQMLGSLPTLYLWLNFCNLYLWPCIVVLIFLAFQQRAYVAPANLPALL	1654
QY	1621	LLLLLYGHSITPLMYPASFPFSPVSTA YVLTCTINLFTGINGSMATFVLELPSDQKQEV	1680
DB	1655	LLLLLYGHSITPLMYPASFPFSPVSTA YVLTCTINLFTGINGSMATFVLELPSDQKQEV	1714
QY	1681	SRILKQVFLIPPHFCLGRGLDMVRNOAWADAFERLGRDQFQSPURWEVGNLWMTQ	1740
DB	1715	SRILKQVFLIPPHFCLGRGLDMVRNOAWADAFERLGRDQFQSPURWEVGNLWMTQ	1774
QY	1741	GPLFLLTLLQLHRSQLLPQPRVRSPLLGBEDEDAVARERVRVQATGQGDVLRNLTK	1800
DB	1775	GPLFLLTLLQLHRSQLLPQPRVRSPLLGBEDEDAVARERVRVQATGQGDVLRNLTK	1834
QY	1801	VYRQRMVPAVDRCLGIPPGECFGLLVNGAGKSTTFRMVTGDTLASGEAVLAGHSVAR	1860
DB	1835	VYRQRMVPAVDRCLGIPPGECFGLLVNGAGKSTTFRMVTGDTLASGEAVLAGHSVAR	1894
QY	1861	EPSSAHLNMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVQATAGSLARLGLSWAD	1920
DB	1895	EPSSAHLNMGYCPQSDAIFELLTGREHLELLARLGRVPEAOVQATAGSLARLGLSWAD	1954
QY	1921	RPAGTYSGNKRLATATALVGDPAVVPDPTTGMDSARRFLWNSLLAVREGSVML	1980
DB	1955	RPAGTYSGNKRLATATALVGDPAVVPDPTTGMDSARRFLWNSLLAVREGSVML	2014
QY	1981	TSHSMECEALCSRLA INMNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAFVAA	2040
DB	2015	TSHSMECEALCSRLA INMNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAFAFVAA	2074


```
QY 1201 QSGSPDAVGRVQGWALTRQOQALLKRPILLARRSRGLPAQIVLPALFVGLALVPSLIV 1260
Db 1201 QSGSPDAVGRVQGWALTRQOQALLKRPILLARRSRGLPAQIVLPALFVGLALVPSLIV 1260
QY 1261 PPGHYPALRLSPMTYGAQVSPSEADPGDGRARLEALLQDAGLEPPVQVHSHRFS 1320
Db 1261 PPGHYPALRLSPMTYGAQVSPSEADPGDGRARLEALLQDAGLEPPVQVHSHRFS 1320
QY 1321 PEVPAEAVKVLASGNWTPSPSPACOSOPGARRLLPCDPAAGGPPPPQAVTGSQEVQ 1380
Db 1321 PEVPAEAVKVLASGNWTPSPSPACOSOPGARRLLPCDPAAGGPPPPQAVTGSQEVQ 1380
QY 1381 NLTRNLSDPLVKTYPRLVKQGLTKKWNVEVYSGSLGGRDPGLPSQQLGSRVSEELW 1440
Db 1381 NLTRNLSDPLVKTYPRLVKQGLTKKWNVEVYSGSLGGRDPGLPSQQLGSRVSEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTAWAHSLODQSLKTFWNNKGWHSWAFVNRASNAIIRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSLODQSLKTFWNNKGWHSWAFVNRASNAIIRAHLP 1500
QY 1501 PGPARRHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLEE 1560
Db 1501 PGPARRHAHSITTLNHPNLTKQOLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLEE 1560
QY 1561 RVTRAKHLQMGSLSPLYWLGNFMDMCNLYLPACIVVLIPLAFOORAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMGSLSPLYWLGNFMDMCNLYLPACIVVLIPLAFOORAYVAPANLPALL 1620
QY 1621 LLLLLYGNISITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLQV 1680
Db 1621 LLLLLYGNISITPLMYPASFFSPSTAYVVLTCINLFIGINGSMATFVLELFSQDKLQV 1680
QY 1681 SRLKQVFLIPHFCLGRGLIDMVRNOAMADAFERLGDQFQSPLRWEVVGKLLAMVQ 1740
Db 1681 SRLKQVFLIPHFCLGRGLIDMVRNOAMADAFERLGDQFQSPLRWEVVGKLLAMVQ 1740
QY 1741 GFLPLFLTLQLHRSOLLQPRVRSPLPLGEDEDEVARERERVQATGQDVLVLRNLTK 1800
Db 1741 GFLPLFLTLQLHRSOLLQPRVRSPLPLGEDEDEVARERERVQATGQDVLVLRNLTK 1800
QY 1801 VTRGQMPAVDRCLGIPGECFGLLVGNAGKSTFRMVTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VTRGQMPAVDRCLGIPGECFGLLVGNAGKSTFRMVTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAHLGMYCQSDAIFELTGREHLELLARLRCVPEAOVAOTAGSGLARLGSWAD 1920
Db 1861 EPSAHLGMYCQSDAIFELTGREHLELLARLRCVPEAOVAOTAGSGLARLGSWAD 1920
QY 1921 RPAGTYSNGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLWNSLIAVREGSVML 1980
Db 1921 RPAGTYSNGNKRKLATALVGDPAVVFDEPTTGMDPSARRFLWNSLIAVREGSVML 1980
QY 1981 TSHSMEECALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTTLTRVPAARSQPAAFVAA 2040
Db 1981 TSHSMEECALCSRLAIMVNGFRCLGSPQHLKGRFAAGHTTLTRVPAARSQPAAFVAA 2040
QY 2041 EPPGSELRAHGRRLRFQIPGRCALARVFGELAVHGAHGVDFSVSQTMLREVFVLYF 2100
Db 2041 EPPGSELRAHGRRLRFQIPGRCALARVFGELAVHGAHGVDFSVSQTMLREVFVLYF 2100
QY 2101 SKDQKDEDETERQKAGVGDVPAPGLQHPKRVYSQFLDDPSTAEVTL 2146
Db 2101 SKDQKDEDETERQKAGVGDVPAPGLQHPKRVYSQFLDDPSTAEVTL 2146
```

RESULT 5
ID ABU08466 standard; protein; 2146 AA.
AC ABU08466;
XX
DT 18-JUN-2003 (first entry)

```
XX Human ABCA-SSN protein.  
DE Human; ATP-binding cassette transporter protein A7; ABC transporter;  
KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;  
KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;  
KW immunomodulator; immunosuppressive; antiinflammatory; ABCA-SSN;  
KW antiarteriosclerotic.  
XX Homo sapiens.  
XX WO2003010315-A1.  
XX 06-FEB-2003.  
XX 24-JUL-2002; 2002WO-JP007487.  
XX 25-JUL-2001; 2001JP-00224176.  
XX 06-DEC-2001; 2001JP-00372530.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX (KAZU-) KAZUSA DNA RES INST FOUND.  
XX Ueda K, Nakagawa S, Nagase T;  
XX WPI; 2003-239444/23.  
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in  
PT the immune system, applicable in diagnosis of and screening drugs for  
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.  
XX  
PS Claim 18; Page 164-174; 183pp; Japanese.  
XX  
CC The present invention relates to the isolation of human ATP-binding  
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the  
CC polynucleotide sequences encoding them. The protein is applicable in the  
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's  
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.  
CC It may also be used in a method for screening ABCA-SSN inhibitors. The  
CC present sequence represents human ABCA-SSN protein  
XX  
SQ Sequence 2146 AA;  
Query Match 99.9%; Score 11128; DB 6; Length 2146;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MAFWTQLMLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60  
Db 1 MAFWTQLMLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNKP 60  
QY 61 LPSAGTVPWLOGLICNVNNTCPQPTPGEEPGRLSNFNDLSVRLADARTVLGGASAH 120  
Db 61 LPSAGTVPWLOGLICNVNNTCPQPTPGEEPGRLSNFNDLSVRLADARTVLGGASAH 120  
QY 121 TLVAGGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQAEPL 180  
Db 121 TLVAGGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLAQAEPL 180  
QY 181 HSLLEAAEDLAQELLALRSVLVELRALLQRPCTSGPFLLSALCSVRGSPSTVGP 240  
Db 181 HSLLEAAEDLAQELLALRSVLVELRALLQRPCTSGPFLLSALCSVRGSPSTVGP 240  
QY 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRRLKPLIIGKLLFAD 300  
Db 241 YEASDLMELVQEPESALPDSLSLSPACSELIGALDSHPLSRLWRRLKPLIIGKLLFAD 300  
QY 301 TPTFKLMAQVNRTEELTLRDVREVWEMLGPRIFTFMNDSNVAMLRLLQMODQGR 360  
Db 301 TPTFKLMAQVNRTEELTLRDVREVWEMLGPRIFTFMNDSNVAMLRLLQMODQGR 360  
QY 361 QPRGGRDHMEALRSFLDPGSGVSWQDAHADVGLVGTGLGRVTCLSLDKLEAAPSEAA 420
```

Db 361 QPRGCRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTECLSLDKLEAA PSEAA 420
Qy 421 LVSRALQLLAHREFWAGVVFGLPBDSSDPTHEPFDLGPQHVRIKIRMDIDVTRTNKIR 480
Db 421 LVSRALQLLAHREFWAGVVFGLPBDSSDPTHEPFDLGPQHVRIKIRMDIDVTRTNKIR 480
Qy 481 DRFDWGPADPLTDLRYVWGGFFVYLODLVERAAVRVLSGANPRAGLYLOMPYPCYVDD 540
Db 481 DRFDWGPADPLTDLRYVWGGFFVYLODLVERAAVRVLSGANPRAGLYLOMPYPCYVDD 540
Qy 541 VFLRVLSRSPLFTLAWIYSVTITVRAVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
Db 541 VFLRVLSRSPLFTLAWIYSVTITVRAVREKETRLDRTRAMGLSRAVLWLGWFLSCL 600
Qy 601 GPFLLSAALVLVLUKGDILPYSHPGVVFPLAFAVAVTQSFLLSAFFSRANLAAACG 660
Db 601 GPFLLSAALVLVLUKGDILPYSHPGVVFPLAFAVAVTQSFLLSAFFSRANLAAACG 660
Qy 661 GLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSVAFSGCESLALLEBEGGAQWENVG 720
Db 661 GLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSVAFSGCESLALLEBEGGAQWENVG 720
Qy 721 TRPTADVFSLAQVSGLLLDAAALXGLATWYLEAVCPGOYGIPEPWNPFRRSYWCGPRPP 780
Db 721 TRPTADVFSLAQVSGLLLDAAALXGLATWYLEAVCPGOYGIPEPWNPFRRSYWCGPRPP 780
Qy 781 KSPAPCPTLPDKVLYVEAPPGSLPGSVRSLEKRPFGSPQALRGLSLDPYQGHITAF 840
Db 781 KSPAPCPTLPDKVLYVEAPPGSLPGSVRSLEKRPFGSPQALRGLSLDPYQGHITAF 840
Qy 841 GHNGAGKTTLSILSGFLPPSGGSAFILGHDRVSSMAAIRPHLGVCPOYNVLPDMLTVDE 900
Db 841 GHNGAGKTTLSILSGFLPPSGGSAFILGHDRVSSMAAIRPHLGVCPOYNVLPDMLTVDE 900
Qy 901 HWFYGRILKGLSAAVGPEDRLQDVLVSKQSVQTRHLSSGMOKLSVAIAFVGGSQV 960
Db 901 HWFYGRILKGLSAAVGPEDRLQDVLVSKQSVQTRHLSSGMOKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGILWELLKYREGRTLILSTHLDLAEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGILWELLKYREGRTLILSTHLDLAEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRHLSGGYLLTVKARLPLTNEKADTMEGSDVTRQKNGSQSRRVGTPOLLA 1080
Db 1021 SPLFLRHLSGGYLLTVKARLPLTNEKADTMEGSDVTRQKNGSQSRRVGTPOLLA 1080
Qy 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRIAELRLTGYSIDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLPYTGADHGSFATLFRDLTRIAELRLTGYSIDTSLE 1140
Qy 1141 EIFLKVVEECAADTDMEDGSCQHLCGTGIAGLDVTLRKMPPOBTALENGEPAGSAPETD 1200
Db 1141 EIFLKVVEECAADTDMEDGSCQHLCGTGIAGLDVTLRKMPPOBTALENGEPAGSAPETD 1200
Qy 1201 QGSGPDVAVRVOGWALTRQQLQALLKRLFLIARSRGLFAQIVLPALFVGLALVFSILV 1260
Db 1201 QGSGPDVAVRVOGWALTRQQLQALLKRLFLIARSRGLFAQIVLPALFVGLALVFSILV 1260
Qy 1261 PPFQHYPALRLSPMYGAQVSVFFSDEAPGDPGRARLLEALLQAEAGLEPPVQVHSSHRFSA 1320
Db 1261 PPFQHYPALRLSPMYGAQVSVFFSDEAPGDPGRARLLEALLQAEAGLEPPVQVHSSHRFSA 1320
Qy 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSBEVQ 1380
Db 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARLLPDCPAAAGPPPPQAVTSGSBEVQ 1380
Qy 1381 NLTCGNLSDFLVKTYPRLVROGLTKKWVNEVRVCGFSGLGRDPLSGQELGRSVEELW 1440
Db 1381 NLTCGNLSDFLVKTYPRLVROGLTKKWVNEVRVCGFSGLGRDPLSGQELGRSVEELW 1440
Qy 1441 ALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSMAFVFNRRASNAILRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSMAFVFNRRASNAILRAHLP 1500

Qy 1501 PGPARRHAHSITTLNHPNLTKQOLSEAAALMASSVDVLVSIQVVFAMSVFASFTVLILIEE 1560
Db 1501 PGPARRHAHSITTLNHPNLTKQOLFEAALMASSVDVLVSIQVVFAMSVFASFTVLILIEE 1560
Qy 1561 RVTRAKHLQMLGGLSPTYLWLGNFMDMCMNYLPACIVVLIIFLAFOQRAVYVAPANLPALL 1620
Db 1561 RVTRAKHLQMLGGLSPTYLWLGNFMDMCMNYLPACIVVLIIFLAFOQRAVYVAPANLPALL 1620
Qy 1621 LLLLXYGWSITPLMYPASFFSVSTAYVVLVTCINLFIGINGSMATFVLELFSQKLOEV 1680
Db 1621 LLLLXYGWSITPLMYPASFFSVSTAYVVLVTCINLFIGINGSMATFVLELFSQKLOEV 1680
Qy 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDROFQSPRLRVEVVGKLLAMVIQ 1740
Db 1681 SRLKQVFLIFPHFCLGRGLIDMVRNQAMADAPERLGDROFQSPRLRVEVVGKLLAMVIQ 1740
Qy 1741 GPLFLFTLLQHRSQLLPQPRVRSPLPLGEEDEEDVARERERVVOGATQGDVLVRLNLT 1800
Db 1741 GPLFLFTLLQHRSQLLPQPRVRSPLPLGEEDEEDVARERERVVOGATQGDVLVRLNLT 1800
Qy 1801 VYRGORMPAVDRLCLGIPPGCECFGLLGVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRLCLGIPPGCECFGLLGVNGAGKTSITFRMVTGDTLASRGEAVLAGHSVAR 1860
Qy 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1920
Db 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGVPEAQVAQTAGSLARLGLSWAD 1920
Qy 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
Db 1921 RPAGTYSGGNKRKLATALVGDPAVVFLDEPTTGMDDPSARRFLMNSLLAVVREGSVML 1980
Qy 1981 TSHSMECEALCRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAASQPAFAA 2040
Db 1981 TSHSMECEALCRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAAASQPAFAA 2040
Qy 2041 EFPQSELREAHGRLRFPQPGRCALARVFGELAVHGAEGHVEDFSVSTMLLEEVLYF 2100
Db 2041 EFPQSELREAHGRLRFPQPGRCALARVFGELAVHGAEGHVEDFSVSTMLLEEVLYF 2100
Qy 2101 SKQQKDEDETEQKEAGVGVDPAFQHPKRVVSQFLDDPSTAETVL 2146
Db 2101 SKQQKDEDETEQKEAGVGVDPAFQHPKRVVSQFLDDPSTAETVL 2146

RESULT 6

AAU09174 ID AAU09174 standard; protein; 2144 AA.
XX AC AAU09174;
XX DT 09-APR-2002 (first entry)
XX DE Human transporter molecule, MTP-1.
XX KW Human; transporter molecule; MTP-1; cytotostatic; Nootropic; HIV; Neuroprotective; Antiparkinsonian; Anticonvulsant; Antianaemic; protein; Antidiabetic; Antiarteriosclerotic; Anti-human immunodeficiency virus; Antiarthritic; Immunosuppressive; Antiasthmatic; Tuberculostatic; Antiulcer; Antimanic; Tranquilliser; Vasotropic; fertility disorder; Transporter-associated disorder; haematopoietic disorder; anxiety; leucocytic disorder; lipid metabolism; abnormal vascularisation; immunological disorder; inflammatory disease; neurological disorder; obsessive-compulsive disorder; cardiac-related disorder; hormonal disorder; reproductive disorder.
XX OS Homo sapiens.
XX FN WO200187978-A2.
XX PD 22-NOV-2001.
XX

PF 14-MAY-2001; 2001WO-US015533.
XX
PR 12-MAY-2000; 2000US-0204211P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Glucksmann M, Curtis RAJ;
XX
DR WPI; 2002-082985/11.
XX N-PSDB; AAS19207.
DR
XX New membrane transport protein and polynucleotides, useful for diagnosing
PT and treating transport protein related disorders e.g. cancer, restenosis,
PT asthma and Alzheimer's disease and to identify modulators of therapeutic
PT use.
XX
PS Claim 12; Fig 1; 141pp; English.
XX
CC The invention relates to an isolated membrane transporter protein-1 (MTP-
CC 1) (I). (I) is useful for identifying a compound which modulates the
CC activity of (I). The method comprises contacting (I) or cell expressing
CC (I) with a test compound and determining whether (I) binds to the test
CC compound or determining the effect of the compound on the activity or
CC expression of (I). The identified compound is useful in treatment and
CC diagnosis of a subject having disorders characterised by aberrant or
CC unwanted MTP-1 protein or nucleic acid expression or activity, where
CC transporter-associated disorders include haematopoietic disorders,
CC leukocytic disorders, disorders related to lipid metabolism, disorders
CC involving abnormal vascularisation, immunological disorders, inflammatory
CC diseases, neurological disorders, anxiety disorders, obsessive-compulsive
CC disorders, cardiac-related disorders. Disorders also include cellular
CC proliferation, growth, differentiation, hormonal disorders and
CC reproductive or fertility disorders. The present sequence represents the
CC amino acid sequence of human transporter molecule, MTP-1
XX
SQ Sequence 2144 AA;

Query Match 99.7%; Score 11107; DB 5; Length 2144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 MAFWTQMLLWKNFMYRRQVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 60
DB 1 MAFWTQMLLWKNFMYRRQVQLLVLLWPLFLFFILVAVRSHHPLEHHECHFPNKP 58
QY 61 LPSAGTVPWLOGLICNVNNTCPQPTGEEPRGLSNFNDLSVSRLLADARTVLGASAH 120
DB 59 LPSAGTVPWLOGLICNVNNTCPQPTGEEPRGLSNFNDLSVSRLLADARTVLGASAH 118
QY 121 TLAGLGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLALGOAEP 180
DB 119 TLAGLGKLIATLRAARSTAOPTKQSPLEPPMLDVAELLTSLRTESLGLALGOAEP 178
QY 181 HSLLEAEADLAQELLALSLVELRALLQRPGTSGPSELLSEALCSVRGSPSTVGPSLNW 240
DB 179 HSLLEAEADLAQELLALSLVELRALLQRPGTSGPSELLSEALCSVRGSPSTVGPSLNW 238
QY 241 YEASDLMELVGPEPSALPDSLSLSPACSELICGALDHPILSRLLWRLKPLILGKLLFAPD 300
DB 239 YEASDLMELVGPEPSALPDSLSLSPACSELICGALDHPILSRLLWRLKPLILGKLLFAPD 298
QY 301 TPFTTKLMAQVNRTEETLLRDREVWMLGPRIFTFMNDSSNVAMLQRLQWDEGRR 360
DB 299 TPFTTKLMAQVNRTEETLLRDREVWMLGPRIFTFMNDSSNVAMLQRLQWDEGRR 358
QY 361 QPRGGRDHMEALRSLDPGSGGYQWQDAHDVGLVGTGLRVTECLSLDKLEAPSEAA 420
DB 359 QPRGGRDHMEALRSLDPGSGGYQWQDAHDVGLVGTGLRVTECLSLDKLEAPSEAA 418
QY 421 LVSRALQLLAHRFWAGVVFVLPEDSSDPTTEHTPDLPFGHVRIKRMIDIVVTRTNKIR 480
DB 419 LVSRALQLLAHRFWAGVVFVLPEDSSDPTTEHTPDLPFGHVRIKRMIDIVVTRTNKIR 478

QY 481 DRFWDGPAADPLTDLRYVMGGFVYLODLVERAAVRVLSGANPRAGLYLQMPYCYVDD 540
DB 479 DRFWDGPAADPLTDLRYVMGGFVYLODLVERAAVRVLSGANPRAGLYLQMPYCYVDD 538
QY 541 VFLRVLSRLPLFLTLAWIYSVTTLTKAVVREKETRLDRDTWRAMGLSRAVLWGLWFLSCL 600
DB 539 VFLRVLSRLPLFLTLAWIYSVTTLTKAVVREKETRLDRDTWRAMGLSRAVLWGLWFLSCL 598
QY 601 GPFLLSAALLVLVLKIGDILPYSHPGVVFLFLAFAVATVTQSFLLSAFFSRANLAAACG 660
DB 599 GPFLLSAALLVLVLKIGDILPYSHPGVVFLFLAFAVATVTQSFLLSAFFSRANLAAACG 658
QY 661 GLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSPVAFGCGCESLALLBEQEGGAQWHNVG 720
DB 659 GLAYFSLYLPVLCVAMRDRLPAGGRVAASLLSPVAFGCGCESLALLBEQEGGAQWHNVG 718
QY 721 TRPTADVESLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPWNFFRPSYWCGRPP 780
DB 719 TRPTADVESLAQVSGLLLLDAAALYGLATWYLEAVCPQYGIPEPWNFFRPSYWCGRPP 778
QY 781 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPALRGLSLDFYQGHITAPL 840
DB 779 KSPAPCTPLDPKVLVEAPPLSGVSVRSLEKRFPGSPALRGLSLDFYQGHITAPL 838
QY 841 GHNGAGKTTTLLSGLPFPSSGSAF ILGHSDVRSSMAAIRPHLGVCPOYNVLFDMLTVD 900
DB 839 GHNGAGKTTTLLSGLPFPSSGSAF ILGHSDVRSSMAAIRPHLGVCPOYNVLFDMLTVD 898
QY 901 HWFVGRUKGLSAAVGPEDRLQDVLVSKQSVQTRHLSGGMORKLSVAIFVGGSQV 960
DB 899 HWFVGRUKGLSAAVGPEDRLQDVLVSKQSVQTRHLSGGMORKLSVAIFVGGSQV 958
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLDEALLGDRVAVVAGGLCCCG 1020
DB 959 VILDEPTAGVDPASRRGIWELLKYREGRTILSTHLDLDEALLGDRVAVVAGGLCCCG 1018
QY 1021 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTROEKNKGQSGSRVGPQLLA 1080
DB 1019 SPLFLRRHLGSGYYLTLVKARLPLTTNEKADTMEGSVDTROEKNKGQSGSRVGPQLLA 1078
QY 1081 LVQHWVPGARLVEBELPHELVLVLPYTGADHSGFATLPRELDTRLAELRLTYGIDSTSL 1140
DB 1079 LVQHWVPGARLVEBELPHELVLVLPYTGADHSGFATLPRELDTRLAELRLTYGIDSTSL 1138
QY 1141 EIFLKVVEECAAADTMEDGSCGQHLCTGIAGLDVTTLKMPPOETALENGPAGAPETD 1200
DB 1139 EIFLKVVEECAAADTMEDGSCGQHLCTGIAGLDVTTLKMPPOETALENGPAGAPETD 1198
QY 1201 QGSGPDVAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1260
DB 1199 QGSGPDVAVGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLPALFVGLALVFSLIV 1258
QY 1261 PPGHYPALRLSPWTYGAQVSFFSEDAPGDPGRARLLLEALLQEALEBPPVQHSRHSFA 1320
DB 1259 PPGHYPALRLSPWTYGAQVSFFSEDAPGDPGRARLLLEALLQEALEBPPVQHSRHSFA 1318
QY 1321 PEVPAEAKVLASGNWTPEPSACOCOPCARLLPDCPAAAGPPPPQAVTSGGEVVQ 1380
DB 1319 PEVPAEAKVLASGNWTPEPSACOCOPCARLLPDCPAAAGPPPPQAVTSGGEVVQ 1378
QY 1381 NLTRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGFSLGGRDPGLPSQBELGRSVEBLW 1440
DB 1379 NLTRNLSDFLVKTYPRLVROGLTKKWNNEVRYGGFSLGGRDPGLPSQBELGRSVEBLW 1438
QY 1441 ALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIFWNKNGHSMVAFVNRASNAILRAHLP 1500
DB 1439 ALLSPLPGGALDRVLKNTAWAHSIDAQDSLKIFWNKNGHSMVAFVNRASNAILRAHLP 1498
QY 1501 PGPARHAHSITTLNHNPLNLTKEQLSEALMASSVDVLVSIQVVFAMSFPVPSFTLVLTBE 1560
DB 1499 PGPARHAHSITTLNHNPLNLTKEQLSEALMASSVDVLVSIQVVFAMSFPVPSFTLVLTBE 1558
QY 1561 RVTRAKHLQMGGLSPTLYWLGNFWDNMCNLYLPACIVVLIPLAFOQRAYVAPANLPALL 1620

Db 1559 RVTRAKHLQLMGGLSPTLYWLNFGNFWLMDNCNVLPACIVVLIFFLAQQRAYVAPANLPALL 1618
Qy 1621 LLLLLYKWSITPLMYPASFFFSVSTAVVLTCTNINLFIGINGSMATFVLELFSQKLOEV 1680
Db 1619 LLLLLYKWSITPLMYPASFFFSVSTAVVLTCTNINLFIGINGSMATFVLELFSQKLOEV 1678
Qy 1681 SRIILKQVFLIEPHCLGRLGLDMVRNQAMADAFERLGDROFQSPRLRHEVVKNNLLAMVIO 1740
Db 1679 SRIILKQVFLIEPHCLGRLGLDMVRNQAMADAFERLGDROFQSPRLRHEVVKNNLLAMVIO 1738
Qy 1741 GPFLLFTLLLOHRSQLLPQPRVRSPLPLGGEDEDAVARERVRVQGATQGDVLRNLTK 1800
Db 1739 GPFLLFTLLLOHRSQLLPQPRVRSPLPLGGEDEDAVARERVRVQGATQGDVLRNLTK 1798
Qy 1801 VYRGQMPAVDRCLGIPPGCFGLGVNGAGKTSFTRMVGTDLASRGSAVLAGHSVAR 1860
Db 1799 VYRGQMPAVDRCLGIPPGCFGLGVNGAGKTSFTRMVGTDLASRGSAVLAGHSVAR 1858
Qy 1861 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSWYAD 1920
Db 1859 EPSAAHLISMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSWYAD 1918
Qy 1921 RPACTYSGGNKRLATLALVGDPAVVFLEDPPTGMDPSARRFLWNSLLAVVREGRSVML 1980
Db 1919 RPACTYSGGNKRLATLALVGDPAVVFLEDPPTGMDPSARRFLWNSLLAVVREGRSVML 1978
Qy 1981 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAPFAVA 2040
Db 1979 TSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAAGHTLTLRVPAARSQPAAPFAVA 2038
Qy 2041 EFPQSELRHAHGRRLRQLPPGGRCALARVFGELAVHGAHGVDFSVQTMLEEVFLYF 2100
Db 2039 EFPQSELRHAHGRRLRQLPPGGRCALARVFGELAVHGAHGVDFSVQTMLEEVFLYF 2098
Qy 2101 SKDQKDEDEEOKAEAGVVDPAFGLQHPKRVSOFLDDPSTAEIVL 2146
Db 2099 SKDQKDEDEEOKAEAGVVDPAFGLQHPKRVSOFLDDPSTAEIVL 2144

RESULT 7

ADD37429
ID ADD37429 standard; protein; 2144 AA.

XX AC ADD37429;

XX DT 15-JAN-2004 (first entry)

XX DE Human transporter MTP-1.

XX KW Human; transporter; cytostatic; anorectic; antidiabetic; anticonvulsant;
KW gene therapy; PGC-1 associated disorder; liver tumour; obesity; epilepsy;
KW diabetes.

XX OS Homo sapiens.

XX XN US2003143675-A1.

XX PD 31-JUL-2003.

XX PF 22-MAY-2002; 2002US-00154419.

XX PR 12-MAY-2000; 2000US-0204211P.

PR 29-JUN-2000; 2000US-0215376P.

PR 31-JUL-2000; 2000US-0221769P.

PR 19-SEP-2000; 2000US-0233790P.

PR 25-SEP-2000; 2000US-0235107P.

PR 05-OCT-2000; 2000US-0238336P.

PR 15-NOV-2000; 2000US-0248364P.

PR 15-DEC-2000; 2000US-0256240P.

PR 18-DEC-2000; 2000US-0256588P.

PR 21-DEC-2000; 2000US-0258028P.

PR 22-JAN-2001; 2001US-0263169P.
PR 14-MAY-2001; 2001US-00858194.
PR 29-JUN-2001; 2001US-00895811.
PR 31-JUL-2001; 2001US-00919781.
PR 19-SEP-2001; 2001US-00957664.
PR 25-SEP-2001; 2001US-00964295.
PR 05-OCT-2001; 2001US-00972724.
PR 14-NOV-2001; 2001US-00002769.
PR 17-DEC-2001; 2001US-00024623.
PR 22-JAN-2002; 2002US-00055025.
XX (MILL-) MILLENNIUM PHARM INC.

XX PA Curtis RAJ, Glucksmann MA, Meyers RE;

XX PI WPI; 2003-851783/79.

XX DR N-PSDB; ADD37428, ADD37430.

XX PT New isolated nucleic acid, useful for preparing a composition for
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
PT or diabetes.

XX PS Claim 11; SEQ ID NO 2; 663pp; English.

XX CC The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumours,
CC obesity, epilepsy or diabetes. The present sequence represents a novel
CC human transporter protein.

XX SQ Sequence 2144 AA;

Query Match 99.7%; Score 11107; DB 7; Length 2144;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2141; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 MAFWTQLMLLWKNFMVRRRQPVQLLVLELWPLFLFFILVAVRHSHPPLHCHFPNKP 60

Db 1 MAFWTQLMLLWKNFMVRRRQPVQLLVLELWPLFLFFILVAVRHSHPPLHCHFPNKP 58

Qy 61 LPSAGTVPMLOGLICNVNNTCFPQLTPGBEPGRLSNFDLSVSRLLADARTVLGGASAH 120

Db 59 LPSAGTVPMLOGLICNVNNTCFPQLTPGBEPGRLSNFDLSVSRLLADARTVLGGASAH 118

Qy 121 TLALGKGLTATLAARSTAQPOPTKOSPPLPPMLDVAELLTSLRTESLGLAQOAEPL 180

Db 119 TLALGKGLTATLAARSTAQPOPTKOSPPLPPMLDVAELLTSLRTESLGLAQOAEPL 178

Qy 181 HSLLAEAAEDLAQELLALRSILVELRALLQRPRTSGPLELSEALCSVRGSPSTVGPGLNW 240

Db 179 HSLLAEAGDLAQELLALRSILVELRALLQRPRTSGPLELSEALCSVRGSPSTVGPGLNW 238

Qy 241 YEASDLMELVQPEPESALPDSSLSLSPACSELIGALDSHPISRLLRRLKPIILKLLFAPD 300

Db 239 YEASDLMELVQPEPESALPDSSLSLSPACSELIGALDSHPISRLLRRLKPIILKLLFAPD 298

Qy 301 TPETRKLMAQVNTFEELTLRDVREVMELGPRIFTEWMDSSNVAMLQRLQMDGERR 360

Db 299 TPETRKLMAQVNTFEELTLRDVREVMELGPRIFTEWMDSSNVAMLQRLQMDGERR 358

Qy 361 QRPFGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLAAPSEAA 420

Db 359 QRPFGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGRVTECLSLDKLAAPSEAA 418

QY 421 LVSRALQLAERHFWAGVVFGLGPESSDPTHTPDLDGEGHVR IKIRMDIDVVTINKIR 480
DB 419 LVSRALQLAERHFWAGVVFGLGPESSDPTHTPDLDGEGHVR IKIRMDIDVVTINKIR 478
QY 481 DRFPDGPAAADPLTDLRYWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYCYDD 540
DB 479 DRFPDGPAAADPLTDLRYWGGFVYLQDLVERAAVRVLSGANPRAGLYLQOMPYCYDD 538
QY 541 VFLRVLSRLPLPLTLAWIYSVTLTKAVVREKETRLRDTMRAMGLSRAVLWGLFSLCL 600
DB 539 VFLRVLSRLPLPLTLAWIYSVTLTKAVVREKETRLRDTMRAMGLSRAVLWGLFSLCL 598
QY 601 GPFLLSAALLVLVLKGLDILPYSHPGVWFLFLAFAFAVATVTSQSFLLSAFFFSRANLAAACG 660
DB 599 GPFLLSAALLVLVLKGLDILPYSHPGVWFLFLAFAFAVATVTSQSFLLSAFFFSRANLAAACG 658
QY 661 GLAYFSLYLPYVLCVAVWRDRLPAGGRVAASLLSPVAFPGCBESLALLBEQGEQAQWNVG 720
DB 659 GLAYFSLYLPYVLCVAVWRDRLPAGGRVAASLLSPVAFPGCBESLALLBEQGEQAQWNVG 718
QY 721 TRPTADVPSLAQVGLLLDAAALYGLATWYLEAVCPGOYGIPEPNWFPFRSYWCGPRPP 780
DB 719 TRPTADVPSLAQVGLLLDAAALYGLATWYLEAVCPGOYGIPEPNWFPFRSYWCGPRPP 778
QY 781 KSPAPCPTPLDPKVLVEEAPPGLSFGSVRSLEKEFPGPSPOPALRGLSLDFYQGHITAPL 840
DB 779 KSPAPCPTPLDPKVLVEEAPPGLSFGSVRSLEKEFPGPSPOPALRGLSLDFYQGHITAPL 838
QY 841 GHNGAGKTTTILSGLPPPGSGSAFIIIGHDVRSSMAAIRPHLGVCPQNVLPDMLTVDE 900
DB 839 GHNGAGKTTTILSGLPPPGSGSAFIIIGHDVRSSMAAIRPHLGVCPQNVLPDMLTVDE 898
QY 901 HWFVGRUKLSAAVVGPEQRLDQVGLVSKQSVQTRHLGGMORKLSVAIAFVGGSOV 960
DB 899 HWFVGRUKLSAAVVGPEQRLDQVGLVSKQSVQTRHLGGMORKLSVAIAFVGGSOV 958
QY 961 VILDEPTAGVDPASRRGIWELLKYREGRTIILSTHLDDEALLGDRVAVVAGGLCCCG 1020
DB 959 VILDEPTAGVDPASRRGIWELLKYREGRTIILSTHLDDEALLGDRVAVVAGGLCCCG 1018
QY 1021 SPLPLRRHLGSGYILTUKARLPLTTNEKADTMEGSDVTRQEKNGQSGSRVGPQLLA 1080
DB 1019 SPLPLRRHLGSGYILTUKARLPLTTNEKADTMEGSDVTRQEKNGQSGSRVGPQLLA 1078
QY 1081 LVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLPRELDTRLAELRLTGYSIDTSLE 1140
DB 1079 LVQHWVPCARLVEELPHELVLVLPYTGADHGSFATLPRELDTRLAELRLTGYSIDTSLE 1138
QY 1141 EIFLKVVECAADTDMEDSGCGHLCTGIAGLDVTLRKMPPEQTALENGBPAGSAPETD 1200
DB 1139 EIFLKVVECAADTDMEDSGCGHLCTGIAGLDVTLRKMPPEQTALENGBPAGSAPETD 1198
QY 1201 QGSGDPAVGRVQGWALTQOOLALLKRLPLARRSRRLPAQIIVLPALFVGLALVFSILV 1260
DB 1199 QGSGDPAVGRVQGWALTQOOLALLKRLPLARRSRRLPAQIIVLPALFVGLALVFSILV 1258
QY 1261 PPGHYPALRLSPMTYGAQVFFSEDAQDPCGRARLLEALLQEAQLEPPVQHSRPSA 1320
DB 1259 PPGHYPALRLSPMTYGAQVFFSEDAQDPCGRARLLEALLQEAQLEPPVQHSRPSA 1318
QY 1321 PEVPAEAVKVLASGNWTPESPSPACQSQPGARRLLPCPAAAGPPPPQAVTSGGEVVQ 1380
DB 1319 PEVPAEAVKVLASGNWTPESPSPACQSQPGARRLLPCPAAAGPPPPQAVTSGGEVVQ 1378
QY 1381 NLTRNLSDFLVKTYPRLVRQGLTKYKWNVVRVGGFSLGGRDPGLPSQELGRSVEELW 1440
DB 1379 NLTRNLSDFLVKTYPRLVRQGLTKYKWNVVRVGGFSLGGRDPGLPSQELGRSVEELW 1438
QY 1441 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSWVAFVNRASNAIILRAHLP 1500
DB 1439 ALLSPLPGALDRVLKNTAWAHSIDAQDSLKIWFNNKGWHSWVAFVNRASNAIILRAHLP 1498

QY 1501 FGPAPHAHSITTLNHNPLNLTKQLSEAAALMASSVDVLVSIYCVVPAMSFVPASFTVLVIEE 1560
DB 1499 FGPAPHAHSITTLNHNPLNLTKQLSEAAALMASSVDVLVSIYCVVPAMSFVPASFTVLVIEE 1558
QY 1561 RVTRAKHQLMGGSLPTLYWLGNFWMCMNYLVPACIVVLIPLAFOORAYVAPANLPALL 1620
DB 1559 RVTRAKHQLMGGSLPTLYWLGNFWMCMNYLVPACIVVLIPLAFOORAYVAPANLPALL 1618
QY 1621 LLLLYLWGSITPLMYPASFFSVPSSTAVVLTCTINLFIGINGSMATFVLELFSDDKLQSV 1680
DB 1619 LLLLYLWGSITPLMYPASFFSVPSSTAVVLTCTINLFIGINGSMATFVLELFSDDKLQSV 1678
QY 1681 SRLKQVFLIPPHFCLGRGLIDMVRNQAMADAPERLGRQFQSPURWEVVGKLLAMVITQ 1740
DB 1679 SRLKQVFLIPPHFCLGRGLIDMVRNQAMADAPERLGRQFQSPURWEVVGKLLAMVITQ 1738
QY 1741 GPLPLLLTLLLOHRSQLLPQPRVRSPLPLGEBEDEVARERERVVQATGQGDVLVRLNLT 1800
DB 1739 GPLPLLLTLLLOHRSQLLPQPRVRSPLPLGEBEDEVARERERVVQATGQGDVLVRLNLT 1798
QY 1801 VYRGQMPAVDRLCLGIPPGECFGLGVNGAGKTSFRMVTGDTLASGEAVLAGHSVAR 1860
DB 1799 VYRGQMPAVDRLCLGIPPGECFGLGVNGAGKTSFRMVTGDTLASGEAVLAGHSVAR 1858
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLARGVPEAQVAQTAGSGLARLGLSWAD 1920
DB 1859 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLARGVPEAQVAQTAGSGLARLGLSWAD 1918
QY 1921 RPAGTYSGGNKRLKATALLVGDPAVPLDPTTGMDSARRFLMNSLLAVVREGSRVNL 1980
DB 1919 RPAGTYSGGNKRLKATALLVGDPAVPLDPTTGMDSARRFLMNSLLAVVREGSRVNL 1978
QY 1981 TSHMEECEALCSRLAIMVNGRFRCLGSPQHLKGRPAAGHTLTLLRVPARSQPAAAFVAA 2040
DB 1979 TSHMEECEALCSRLAIMVNGRFRCLGSPQHLKGRPAAGHTLTLLRVPARSQPAAAFVAA 2038
QY 2041 EFPGSELREAHGGRRLRFPGLPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
DB 2039 EFPGSELREAHGGRRLRFPGLPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2098
QY 2101 SKDGKQEDTBEQKEAGVGVDPAPGLQHPKRVSQPLDDPSTAEVTL 2146
DB 2099 SKDGKQEDTBEQKEAGVGVDPAPGLQHPKRVSQPLDDPSTAEVTL 2144

RESULT 8
ABUS4629
ID ABUS4629 standard; protein; 2059 AA.
XX AC ABUS4629;
XX DT 03-JUN-2003 (first entry)
XX DE Human NOVX polypeptide #88.
XX KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW aortic septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerositis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
XX
XX WO200281498-A2.
XX PD 17-OCT-2002.
XX PF 03-APR-2002; 2002WO-US010780.
XX PR 03-APR-2001; 2001US-0281086P.

PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 06-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294184P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 27-SEP-2001; 2001US-0324802P.
PR 17-OCT-2001; 2001US-0325684P.
PR 14-NOV-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX
FA (CURA-) CURAGEN CORP.
XX
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SK, Stone DJ;
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
XX WPI: 2003-046858/04.
DR N-PSDB; ABX72257.
XX
PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 281-282; 66pp; English.
XX
CC The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberos sclerosi, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABU5452-ABU54647 represent human NOVX polypeptides
CC of the invention
XX
SQ Sequence 2059 AA;
XX
Query Match 95.5%; Score 10639.5; DB 6; Length 2059;

Best Local Similarity 95.9%; Pred. No. 0;
Matches 2058; Conservative 0; Mismatches 1; Indels 87; Gaps 1;
Qy 1 MAFWTQLMLLWKNFMYRRRQPVQLLVELLWPLFLFPILVAVRSHSPLEHHECHFPNKP 60
Db 1 MAFWTQLMLLWKNFMYRRRQPVQLLVELLWPLFLFPILVAVRSHSPLEHHECHFPNKP 60
Qy 61 LPSAGTVPWLOGLICNVNNTCFPOLTPGEEPGRLSNFNDLSVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPWLOGLICNVNNTCFPOLTPGEEPGRLSNFNDLSVSRLLADARTVLGGASAH 120
Qy 121 TLAGLGLKIATLAARSTAOPTKOSPPLPMDVAELTSLRTESLGLALGOAEPL 180
Db 121 TLAGLGLKIATLAARSTAOPTKOSPPLPMDVAELTSLRTESLGLALGOAEPL 180
Qy 181 HSLLEAAEDLAQELLALRSIVELRALIQRPGTSGPLELSEALCSVRGFSSTVGP SLNW 240
Db 181 HSLLEAAEDLAQELLALRSIVELRALIQRPGTSGPLELSEALCSVRGFSSTVGP SLNW 240
Qy 241 YEASDLMELVQEPESALPDSSLS PACSELIGALDSHPLSRLLWRLKPLILGKLLFAPD 300
Db 241 YEASDLMELVQEPESALPDSSLS PACSELIGALDSHPLSRLLWRLKPLILGKLLFAPD 300
Qy 301 TPTRKLMQVNRTEBELTLLRDVREVMGLPRIPTFMNDSSNVAMLQRLLOMDGRR 360
Db 301 TPTRKLMQVNRTEBELTLLRDVREVMGLPRIPTFMNDSSNVAMLQRLLOMDGRR 360
Qy 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAPSEAA 420
Db 361 QPRPGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAPSEAA 420
Qy 421 LVSRALQLLAHHRFWAGVFLGPEDSDPTHEHTDPLGPGHVRIRKIMDDVVTNKRIR 480
Db 421 LVSRALQLLAHHRFWAGVFLGPEDSDPTHEHTDPLGPGHVRIRKIMDDVVTNKRIR 480
Qy 481 DRFWDGCPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAGILQOMPYPVCVDD 540
Db 481 DRFWDGCPAADPLTDLRYVMGGFVYLQDLVERAAVRVLSGANPRAGILQOMPYPVCVDD 540
Qy 541 VFLRVLSRSLPLFLTLAWIYSVTITVKA VVREKTRLRDTMRAMGLSRAVLWLGWFLSCL 600
Db 541 VFLRVLSRSLPLFLTLAWIYSVTITVKA VVREKTRLRDTMRAMGLSRAVLWLGWFLSCL 600
Qy 601 GPFLLSAALLVILKLGDIILPYSHPGVWFLFLAFAVATVTSQFLLSAFSPSRANLAACG 660
Db 601 GPFLLSAALLVILKLGDIILPYSHPGVWFLFLAFAVATVTSQFLLSAFSPSRANLAACG 660
Qy 661 GLAYFSLYLPVILCVAWDRDL PAGGRVAASLLSPVAFPGCESLALLEOEGEQAHNVG 720
Db 661 GLAYFSLYLPVILCVAWDRDL PAGGRVAASLLSPVAFPGCESLALLEOEGEQAHNVG 720
Qy 721 TRPTADVFSUAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPWNPFRRSYWCGPRPP 780
Db 721 TRPTADVFSUAQVSGLLLDAAALYGLATWYLEAVCPGQYGIPEPWNPFRRSYWCGPRPP 780
Qy 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVS VRSLEKRPFGSPQALRGSLDFYQGHITAF 840
Db 781 KSPAPCPTPLDPKVLVEEAPPGLSPGVS VRSLEKRPFGSPQALRGSLDFYQGHITAF 840
Qy 841 GHNGAGKTTTILSILSGLFPFGSGSAFTLGH D VRSMAA I RPHLGVCPQYNVLFDMLTVD 900
Db 841 GHNGAGKTTTILSILSGLFPFGSGSAFTLGH D VRSMAA I RPHLGVCPQYNVLFDMLTVD 900
Qy 901 HVMFYGRKGLSAAVVGPEQDRLLQDVGLSVKSVQVTRHLSGGNQKLSVAIAFVGGSQV 960
Db 901 HVMFYGRKGLSAAVVGPEQDRLLQDVGLSVKSVQVTRHLSGGNQKLSVAIAFVGGSQV 960
Qy 961 VILDEPTAGVDPASRRGIWELLILKYREGRTLI LSTHHLDEAILLGDRAVAVAGRLCCCG 1020
Db 961 VILDEPTAGVDPASRRGIWELLILKYREGRTLI LSTHHLDEAILLGDRAVAVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGSGVYTLVVKARLPLTNEKADTDMEGSVDTROEKKNGSGSRVGTGPOLLA 1080
Db 1021 SPLFLRRHLGSGVYTLVVKARLPLTNEKADTDMEGSVDTROEKKNGSGSRVGTGPOLLA 1080

Db 1021 SPFLRRHLSGYYTLVKARLPLTTNEKADTMEGSDVTRQEKNGSQSRVGTPLQLA 1080
QY 1081 LVQHWVPGARLVEELPHELVLVLVLYTGAHDGSPATLFRDLRLAELRLTGYGSDTSL 1140
Db 1081 LVQHWVPGARLVEELPHELVLVLVLYTGAHDGSPATLFRDLRLAELRLTGYGSDTSL 1140
QY 1141 EFLKVVBECAADTMEDGSCQHLCCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
Db 1141 EFLKVVBECAADTMEDGSCQHLCCTGIAGLDVTLRLKMPPOETALENGEPAGSAPETD 1200
QY 1201 QSGGPDVGRVGVWALTRQOLQALLKRELLARRSRGLFAQIVLPALFVGLALVFLSIV 1260
Db 1201 QSGGPDVGRVGVWALTRQOLQALLKRELLARRSRGLFAQIVLPALFVGLALVFLSIV 1260
QY 1261 PFGHYPALRLSPTMYGAQVSFSEDAQDGPGRARLLEALLQEALEEPVQHSRRFSA 1320
Db 1261 PFGHYPALRLSPTMYGAQVSFSEDAQDGPGRARLLEALLQEALEEPVQHSRRFSA 1320
QY 1321 PEVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVVQ 1380
Db 1321 PEVPAEVAKVLAGSNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVVQ 1380
QY 1381 NLTGRLNSDPLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPGLPSQOELGRSVEELW 1440
Db 1381 NLTGRLNSDPLVKTYPRLVROGLTKKWNVEVRYGFSGLGRDPGLPSQOELGRSVEELW 1440
QY 1441 ALLSLPLGGALDRVLKNTLAWAHSLLDAQDSLKIWFNNKGWHSWAFVNRASNAILRAHLP 1500
Db 1441 ALLSLPLGGALDRVLKNTLAWAHSLLDAQDSLKIWFNNKGWHSWAFVNRASNAILRAHLP 1500
QY 1501 PGPARHAHSITTLNHLPLNTEQLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLE 1560
Db 1501 PGPARHAHSITTLNHLPLNTEQLSEALMASSVDVLVSIQVVFAMSFVPASFTLVLE 1560
QY 1561 RVTRAKHLQMLGSLFTLVNGLNFWDMCNVYLVPACIVLIFLAFQORAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMLGSLFTLVNGLNFWDMCNVYLVPACIVLIFLAFQORAYVAPANLPALL 1620
QY 1621 LLLLLYGSITPLMYPASFPFSPSTAYVVLTCINLFIGINGSMATFVLELFDKQLOEV 1680
Db 1589 -----KLQEV 1593
QY 1681 SRLKQVFLIFPHFCIGRLIDMVRNQAMADAFERLGRDQFQSPLRWEVWGNLLAMVIQ 1740
Db 1594 SRLKQVFLIFPHFCIGRLIDMVRNQAMADAFERLGRDQFQSPLRWEVWGNLLAMVIQ 1653
QY 1741 GFLFLLFTLLQHRSQLLPQPRVRSPLLGEDEDAVARERVRVQATQGDVLRNLTK 1800
Db 1654 GFLFLLFTLLQHRSQLLPQPRVRSPLLGEDEDAVARERVRVQATQGDVLRNLTK 1713
QY 1801 VYRGORMPAVDRLCLIGIPGECFGLLVNGAGKTSTFRWVTGDTLASRGEAVLAGHSVAR 1860
Db 1714 VYRGORMPAVDRLCLIGIPGECFGLLVNGAGKTSTFRWVTGDTLASRGEAVLAGHSVAR 1773
QY 1861 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSWAD 1920
Db 1774 EPSAAHLSMGYCPQSDAIFELLTGREHLELLARLGRVPEAQVAQTAGSLARLGLSWAD 1833
QY 1921 RPAGTYSGNKRLATALALVGDPAVFLDEPTTGMDSARRFLNWSLLAVVREGSRVNL 1980
Db 1834 RPAGTYSGNKRLATALALVGDPAVFLDEPTTGMDSARRFLNWSLLAVVREGSRVNL 1893
QY 1981 TSHSMEECALCSRLAIWNGRFRCLGSPHLLKGRPAAGHTLTLPVPAARSQPAAFVAA 2040
Db 1894 TSHSMEECALCSRLAIWNGRFRCLGSPHLLKGRPAAGHTLTLPVPAARSQPAAFVAA 1953
QY 2041 EPPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2100
Db 1954 EPPGSELREAHGRLRFOLPPGRCALARVFGELAVHGAEGHVEDFSVQTMLEEVFLYF 2013
QY 2101 SKDQKDETEQKAGVGVDPAPGLQHPKRVVSQFLDDPSTAEVTL 2146
Db 2014 SKDQKDETEQKAGVGVDPAPGLQHPKRVVSQFLDDPSTAEVTL 2059

RESULT 9

ABU08464

ID ABU08464 standard; protein; 2008 AA.

XX AC ABU08464;

XX DT 18-JUN-2003 (first entry)

XX DE Amino acid sequence for human ABCA7 splice variant #1.

XX KW Human; ATP-binding cassette transporter protein A7; ABC transporter;
ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
immunomodulator; immunosuppressive; antiinflammatory;
antiarteriosclerotic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Region 1..28 "Given as SEQ ID No:9 and specifically claimed in
Claim 2"

FT FT WO2003010315-A1.

PN PD 06-FEB-2003.

PF PF 24-JUL-2002; 2002WO-JP007487.

PR PR 25-JUL-2001; 2001JP-00224176.

PR PR 06-DEC-2001; 2001JP-00372530.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PA (KAZU-) KAZUSA DNA RES INST FOUND.

XX PI Ueda K, Nakagawa S, Nagase T;

DR DR WPI; 2003-239444/23.

DR DR N-PSDB; ABX95283.

XX PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
the immune system, applicable in diagnosis of and screening drugs for
e.g. autoimmune diseases, Sjogren's syndrome and inflammations.

XX PS Claim 1; Page 106-115; 183pp; Japanese.

XX CC The present invention relates to the isolation of human ATP-binding
cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
polynucleotide sequences encoding them. The protein is applicable in the
diagnosis and screening of drugs for autoimmune diseases, Sjogren's
syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
It may also be used in a method for screening ABCA-SSN inhibitors. The
present sequence represents human ABCA7 splice variant #1

SQ Sequence 2008 AA;

Query Match 92.1%; Score 10264; DB 6; Length 2008;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ESLGLNGQAQCEPLHSLLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELSEALCS 226

Db 29 ESLGLNGQAQCEPLHSLLEAAEDLAQELLALRSVLVELRALLQRPRTSGPLELSEALCS 88

QY 227 VRGSSSTVGPSINWYEASDLMLVQEPESALPDSSLSPPACSELIGALDHPRLWRR 286

Db 89 VRGSSSTVGPSINWYEASDLMLVQEPESALPDSSLSPPACSELIGALDHPRLWRR 148

QY 287 LKPLILGKLLFAPDTPFTRKLMQVNRFTFELLRLDRVREVMGLGPRIFTFMDSSNVA 346

Db 149 LKPLILGKLLFAPDTPFTRKLMQVNRFTFELLRLDRVREVMGLGPRIFTFMDSSNVA 208

QY 347 MLQRLQMDGRRQPPGGRDHMEALRSFLDPCSGGYSMDADHADVCHLVGTIGRVTEC 406
Db 209 MLQRLQMDGRRQPPGGRDHMEALRSFLDPCSGGYSMDADHADVCHLVGTIGRVTEC 268
QY 407 LSLDKLEAAPSEAAALVSRALQLLAHHRFWAGVFLGPESSDPTHEPTPOLGPGHVRKI 466
Db 269 LSLDKLEAAPSEAAALVSRALQLLAHHRFWAGVFLGPESSDPTHEPTPOLGPGHVRKI 328
QY 467 RMDIDVTRTNKIRDRFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526
Db 329 RMDIDVTRTNKIRDRFMDPGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 388
QY 527 LYLOOMPYPYCVDDVDFLVLRSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRMTAMGL 586
Db 389 LYLOOMPYPYCVDDVDFLVLRSRLPLFLTLAWIYSVTLTVKAVVREKETRLDRMTAMGL 448
QY 587 SRAVLWLGWFLSLCLGPFLLSAAALVVLKLGDIILPYSHPGVVFLLAFAFVATVTSQFLL 646
Db 449 SRAVLWLGWFLSLCLGPFLLSAAALVVLKLGDIILPYSHPGVVFLLAFAFVATVTSQFLL 508
QY 647 SAFPSRANLAAACGGLAYFSLPYLVCVWRDLRDPAGGRVAAASLLSPVAFGFCESLAL 706
Db 509 SAFPSRANLAAACGGLAYFSLPYLVCVWRDLRDPAGGRVAAASLLSPVAFGFCESLAL 568
QY 707 LEEQGEQAQHNVTCTRTADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQYQYIPEPWN 766
Db 569 LEEQGEQAQHNVTCTRTADVFSLAQVSGLLILLDAALYGLATWYLEAVCPQYQYIPEPWN 628
QY 767 FPFRRSYWCGPRPKSPAPCTPDLPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 826
Db 629 FPFRRSYWCGPRPKSPAPCTPDLPKVLVEEAPPGLSPGVSVRSLEKRPFGSPQPALRG 688
QY 827 LSLDFYQGHITAFIAGHNGAGKTTTSLISGLFPPSGGSFALGHVDRSSMAAIRPHLGVC 886
Db 689 LSLDFYQGHITAFIAGHNGAGKTTTSLISGLFPPSGGSFALGHVDRSSMAAIRPHLGVC 748
QY 887 PQYNVLFDMLTVDEHWFYGLKLSAANVGPEQDRILLQDVLVSKQSVQTRHLSGMQR 946
Db 749 PQYNVLFDMLTVDEHWFYGLKLSAANVGPEQDRILLQDVLVSKQSVQTRHLSGMQR 808
QY 947 KLSVAIAFVGSSQVVIDEPTAGVDYPASRRGIWELLKLYREGRTLILSTHHLDAELLGD 1006
Db 809 KLSVAIAFVGSSQVVIDEPTAGVDYPASRRGIWELLKLYREGRTLILSTHHLDAELLGD 868
QY 1007 RVAVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTNEKADTMEGSDVDTROBKN 1066
Db 869 RVAVAGRLCCCGSPFLRRHLGSGYYLTLVKARLPLTNEKADTMEGSDVDTROBKN 928
QY 1067 GSQSRVGTQOLLALVQHWVPGARLVEELPHELVLVLPYTGADHSGSFATLFRDLDTLAE 1126
Db 929 GSQSRVGTQOLLALVQHWVPGARLVEELPHELVLVLPYTGADHSGSFATLFRDLDTLAE 988
QY 1127 LRLTGYGIDSLSLEEIFLKVVVECAADTDMEDSCGQHLCTGAGLDVTLRLKMPPOETA 1186
Db 989 LRLTGYGIDSLSLEEIFLKVVVECAADTDMEDSCGQHLCTGAGLDVTLRLKMPPOETA 1048
QY 1187 LENGEPAGSPETDQSGPDVAGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP 1246
Db 1049 LENGEPAGSPETDQSGPDVAGRVQGWALTRQQLQALLKRLFLARRSRRLGFAQIVLP 1108
QY 1247 ALFVGLALVSLIYVPPFGHPYALRLSTMTYGAQVSFFSEADPGDPGARLLEALLQAGL 1306
Db 1109 ALFVGLALVSLIYVPPFGHPYALRLSTMTYGAQVSFFSEADPGDPGARLLEALLQAGL 1168
QY 1307 BEPPVQSHSHRFSAPPEVAEVAKVLASGNMTPESSPACOCQPCARRLLPDCPAAAGGP 1366
Db 1169 BEPPVQSHSHRFSAPPEVAEVAKVLASGNMTPESSPACOCQPCARRLLPDCPAAAGGP 1228
QY 1367 PPOAVTGSSEVVQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGGFSLGGRDPGL 1426
Db 1229 PPOAVTGSSEVVQNLTRNLSDFLVKTYPRLVQGLTKKWNVEVRYGGFSLGGRDPGL 1288

QY 1427 PSQELGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSIDAQDSLKIWFNNKGHWSVAF 1486
Db 1289 PSQELGRSVEELWALLSLPLPGGALDRVLKNLTAWAHSIDAQDSLKIWFNNKGHWSVAF 1348
QY 1487 VNRASNAILRAHLPFGPARHAHSITTLTNHPLNLTKEQSEAAALMASSVDVLVSIYVVFAM 1546
Db 1349 VNRASNAILRAHLPFGPARHAHSITTLTNHPLNLTKEQSEAAALMASSVDVLVSIYVVFAM 1408
QY 1547 SFVPASFTLLVIBERTVTRAKHLQMGGLSPTLIVLGNFLWDMCNLYLPACIVVLIFLAFQ 1606
Db 1409 SFVPASFTLLVIBERTVTRAKHLQMGGLSPTLIVLGNFLWDMCNLYLPACIVVLIFLAFQ 1468
QY 1607 QRAVAPANLPALLLLLLXGWSITPLMYPASFFESVPSTAYVVLTCINLFIGINGSMAT 1666
Db 1469 ORAYVAPANLPALLLLLLXGWSITPLMYPASFFESVPSTAYVVLTCINLFIGINGSMAT 1528
QY 1667 FVLELFSQDLQBSVSRILKQVFLIFPHFCLGRGLIDMVRNQANMADAFERLGDQFQSPLR 1726
Db 1529 FVLELFSQDLQBSVSRILKQVFLIFPHFCLGRGLIDMVRNQANMADAFERLGDQFQSPLR 1588
QY 1727 WEVVGKXLLAMVIOGPFLLFTLLQLHRSQLLPQPRVRSPLPILGEEDEDVARERERVQ 1786
Db 1589 WEVVGKXLLAMVIOGPFLLFTLLQLHRSQLLPQPRVRSPLPILGEEDEDVARERERVQ 1648
QY 1787 ATQGDVLVLRNLTKVYRGQRMPPAVDRLCLGIPPGECFGLGVNGAGKTSFRVMTGDTLA 1846
Db 1649 ATQGDVLVLRNLTKVYRGQRMPPAVDRLCLGIPPGECFGLGVNGAGKTSFRVMTGDTLA 1708
QY 1847 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLRCVPEAQVAQTA 1906
Db 1709 SRGEAVLAGHSVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLARLRCVPEAQVAQTA 1768
QY 1907 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDDPSARRFLWN 1966
Db 1769 GSGLARLGLSWYADRPAGTYSGGNKRLATALAVGDPVAVFLDEPTTGMDDPSARRFLWN 1828
QY 1967 SLIAVVRGREGSVMLTSHSMEECALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLY 2026
Db 1829 SLIAVVRGREGSVMLTSHSMEECALCSRLAIWNGRFRCLGSPQHLKGRFAAGHTLTLY 1888
QY 2027 PAARSQPAAFVAAEFPGSELREAHGRLRFQLPQGRCALARVFGELAVHGAHGVEDF 2086
Db 1889 PAARSQPAAFVAAEFPGSELREAHGRLRFQLPQGRCALARVFGELAVHGAHGVEDF 1948
QY 2087 SVSQTMLEEVFLYFSKQDKDEDETEBQKEAGVGDPAQGLQHPKRVSOQLDDPSTARTVL 2146
Db 1949 SVSQTMLEEVFLYFSKQDKDEDETEBQKEAGVGDPAQGLQHPKRVSOQLDDPSTARTVL 2008
RESULT 10
ABU08465
ID ABU08465 standard; protein; 1993 AA.
XX AC ABU08465;
XX AC ABU08465;
XX AC ABU08465;
DT 18-JUN-2003 (first entry)
XX DE Amino acid sequence for human ABCA7 splice variant #2.
XX DE Human; ATP-binding cassette transporter protein A7; ABC transporter;
XX KW ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
XX KW abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
XX KW immunomodulator; immunosuppressive; antiinflammatory;
XX KW antiarteriosclerotic.
OS Homo sapiens.
XX OS Homo sapiens.
XX OS Homo sapiens.
PN WO2003010315-A1.
XX WO2003010315-A1.
PD 06-FEB-2003.
XX 06-FEB-2003.
PF 24-JUL-2002; 2002WO-JP007487.
XX 24-JUL-2002; 2002WO-JP007487.

PR 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX WPI; 2003-239444/23.
DR N-PSDB; ABX95284.
XX
XX Novel ABC transporter protein, ABCA7 splicing variant, participating in
PT the immune system, applicable in diagnosis of and screening drugs for
PT e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
XX Claim 1; Page 145-154; 183pp; Japanese.
XX
CC The present invention relates to the isolation of human ATP-binding
CC cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
CC polynucleotide sequences encoding them. The protein is applicable in the
CC diagnosis and screening of drugs for autoimmune diseases, Sjogren's
CC syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
CC It may also be used in a method for screening ABCA-SSN inhibitors. The
CC present sequence represents human ABCA7 splice variant #2
XX
SQ Sequence 1993 AA;

Query Match 91.3%; Score 10171.5; DB 6; Length 1993;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

QY 167 ESIGLALGQAQEPFLSHLEAEADLAQELALRSVLVELRALLQRPRTGTSGLLEALCS 226
DB 29 ESIGLALGQAQEPFLSHLEAEADLAQELALRSVLVELRALLQRPRTGTSGLLEALCS 88

QY 227 VRGSPSTVGPSLNWYEASDLMBELVQEPESALPDSLSLSPACSELICALDSHPLSLRWR 286
DB 89 VRGSPSTVGPSLNWYEASDLMBELVQEPESALPDSLSLSPACSELICALDSHPLSLRWR 148

QY 287 LKPLILGKLLFAPDPTFKMAQVNRTPFEELTLARDVREVMELGPRIFTFMDSSNVA 346
DB 149 LKPLILGKLLFAPDPTFKMAQVNRTPFEELTLARDVREVMELGPRIFTFMDSSNVA 208

QY 347 MLQRLQMDQERRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 406
DB 209 MLQRLQMDQERRPRGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLGRVTEC 268

QY 407 LSLDKLEAAPSAAALVSRALQLLASHRFWAGVVFILGPDSSDPTBHTPDGLPGHVRKI 466
DB 269 LSLDKLEAAPSAAALVSRALQLLASHRFWAGVVFILGPDSSDPTBHTPDGLPGHVRKI 328

QY 467 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 526
DB 329 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWGGFVYLQDLVERAAVRVLSGANPRAG 388

QY 527 LYLOQMPYCYVDDVFLRSLPLFLTLAWISVTLTKAVVREKETRLDRDTRAMGL 586
DB 389 LYLOQMPYCYVDDVFLRSLPLFLTLAWISVTLTKAVVREKETRLDRDTRAMGL 448

QY 587 SRAVLWLGWFLSCLGPFLLSAALLVLLKGLDILPYSHPGVVFLLAAPAVATVTSFLL 646
DB 449 SRAVLWLGWFLSCLGPFLLSAALLVLLKGLDILPYSHPGVVFLLAAPAVATVTSFLL 508

QY 647 SAFFSRANLAACGGIAYFSLYLPVLCVWRDRLPAGGRVAASLLSPVAFGCGESLAL 706
DB 509 SAFFSRANLAACGGIAYFSLYLPVLCVWRDRLPAGGRVAASLLSPVAFGCGESLAL 568

QY 707 LEEQEGQAHWNVTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPWN 766
DB 569 LEEQEGQAHWNVTRPTADVFSLAQVSGLLLLDAALYGLATWYLEAVCPQGYGPEPWN 628

QY 767 FPFRRSYWCGPRPPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRG 826
DB

DB 629 FPFRRSYWCGPRPPKSPAPCPTPLDPKVLVEAPPGLSPGVSVRSLEKRFPGSPQALRG 688
QY 827 LSLDFVQGHITAFGLHNGAGKTTLSIISGLPBGSGSAFIIIGHDVRSSMAAIRPHLGYC 886
DB 689 LSLDFVQGHITAFGLHNGAGKTTLSIISGLPBGSGSAFIIIGHDVRSSMAAIRPHLGYC 748
QY 887 POYNVLFDMLTVDDEHWFYGRLLKGLSAAVVGPEQDRLLQDVGVLVSKQSVQTRHLSGCMOR 946
DB 749 POYNVLFDMLTVDDEHWFYGRLLKGLSAAVVGPEQDRLLQDVGVLVSKQSVQTRHLSGCMOR 808
QY 947 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTILSLTHHDEALLGD 1006
DB 809 KLSVAIAFVGGSQVILDEPTAGVDPASRRGIWELLKYREGRTILSLTHHDEALLGD 868
QY 1007 RVAVVAGRLCCCGSPLFLRRHLGSGYYILTVKARLPLTTNEKADTDMEGSVDTROEKN 1066
DB 869 RVAVVAGRLCCCGSPLFLRRHLGSGYYILTVKARLPLTTNEKADTDMEGSVDTROEKN 928
QY 1067 GSQGSRVGTPOQLLALVQHWVPGARLVEBELPHELVLVLPVYTGADHGSFATLFRLEDTRLAE 1126
DB 929 GSQGSRVGTPOQLLALVQHWVPGARLVEBELPHELVLVLPVYTGADHGSFATLFRLEDTRLAE 988
QY 1127 LRLTYGISTSLSEBIFLKVVEECAADTDMEDGSCGQHLCTGCIAGLDVTLRLKMPPOETA 1186
DB 989 LRLTYGISTSLSEBIFLKVVEECAADTDMEDGSCGQHLCTGCIAGLDVTLRLKMPPOETA 1048
QY 1187 LENGEPAGSAPETDOGSGPDVGRVQGWALTQQLQALLKFLFLARSRRLGFAQIVLP 1246
DB 1049 LENGEPAGSAPETDOGSGPDVGRVQGWALTQQLQALLKFLFLARSRRLGFAQIVLP 1108
QY 1247 ALFVGLALVFSILVPPFGHYPALRLSPVWYGAQVSFFSEDPGDPGRARLLLEALLQEAGL 1306
DB 1109 ALFVGLALVFSILVPPFGHYPALRLSPVWYGAQVSFFSEDPGDPGRARLLLEALLQEAGL 1168
QY 1307 BEPPVQSHSRFSAPEVPAEVAKVLAASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1366
DB 1169 BEPPVQSHSRFSAPEVPAEVAKVLAASGNWTPESPSPACQSQPGARRLLPDCPAAAGP 1228
QY 1367 PPQAVTSGGEVQNLGTRNLSDFLVKTYPRLVROGLTKKWVNEVRVYGGFSLGGRDGL 1426
DB 1229 PPQAVTSGGEVQNLGTRNLSDFLVKTYPRLVROGLTKKWVNEVRVYGGFSLGGRDGL 1288
QY 1427 PSGQLGSRVSEBELWALLSPLPGGLADRLVLKNTLTAHSLDAQDSLKIWFNNKGHWSWAF 1486
DB 1289 PSGQLGSRVSEBELWALLSPLPGGLADRLVLKNTLTAHSLDAQDSLKIWFNNKGHWSWAF 1348
QY 1487 VNRSNALILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAM 1546
DB 1349 VNRSNALILRAHLPPGPARHAHSITTLNHPNLNLTKEQLSEAAALMASSVDVLVSIQVVFAM 1408
QY 1547 SFVPASFVTLVLEERVTRAKHLQLMGGISPTLYLGNFLWDMCNLYLPACIVVLIPLAQ 1606
DB 1409 SFVPASFVTLVLEERVTRAKHLQLMGGISPTLYLGNFLWDMCNLYLPACIVVLIPLAQ 1468
QY 1607 QRAYVAPANLALLLLLYGMSITPLMYPASFFFSVPSTAYVLTICNLFIGINGSMAT 1666
DB 1469 QRAYVAPANLALLLLLYGMSITPLMYPASFFFSVPSTAYVLTICNLFIGINGSMAT 1528
QY 1667 FVLELFSQKLOEVSRIKQVFLIPPHFCGLGRGLIDMVNRQAMADA FRLGDRQFQSPUR 1726
DB 1529 FVLELFSQKLOEVSRIKQVFLIPPHFCGLGRGLIDMVNRQAMADA FRLGDRQFQSPUR 1588
QY 1727 WEVWKNLAWIIOGPLELFTLLLOHRSOLLPOPRVRSPLPLGGEDEDAVARERVVQG 1786
DB 1589 WEVWKNLAWIIOGPLELFTLLLOHRSOLLPOPRVRSPLPLGGEDEDAVARERVVQG 1648
QY 1787 ATQGDVLVLRNLTKVYRGORMPAVDRCLGIPPGECFGLLVNGAGKSTFRMVTGDTILA 1846
DB 1649 ATQGDVLVLRNLTKVYRGORMPAVDRCLGIPPGECFGLLVNGAGKSTFRMVTGDTILA 1708
QY 1847 SRGAVALAGHSAVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLGRVPEAQVOTA 1906
DB 1709 SRGAVALAGHSAVAREPSAAHLSMGYCPQSDAIFELLTGREHLELLIARLGRVPEAQVOTA 1768

Qy 1907 GSGIARLGSLWYADRPAGTYSGGNKRKRLATLALVGPVAVFLDEPTTGMDPSARRFLWN 1966
Db 1769 GSGIARLGSLWYADRPAGTYSGGNKRKRLATLALVGPVAVFLDEPTTGMDPSARRFLWN 1828
Qy 1967 SLLAVREGSVMLTSHSMECEALCSRLAIWNGRFRCLGSPQHLKGRPAAGHTLTLRV 2026
Db 1829 SLLAVREGSVMLTSHSMECEALCSRLAIWNGRFRCLGSPQHLKG----- 1876
Qy 2027 PAARSQPAAFVAAEFPGSELREAHGGRRLRQLPFGGRCCALARVFGELAVHGAHGVEDF 2086
Db 1877 ---RSQPAAFVAAEFPGSELREAHGGRRLRQLPFGGRCCALARVFGELAVHGAHGVEDF 1933
Qy 2087 SVSQTMLLEEVLYFSKQKQKDETEBQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 2146
Db 1934 SVSQTMLLEEVLYFSKQKQKDETEBQKEAGVGVDPAPGLQHPKRVSQFLDDPSTAETVL 1993
RESULT 11
AAU04484
ID AAU04484 standard; protein; 1873 AA.
AC AAU04484;
XX XX
XX 26-SEP-2001 (first entry)
XX Human PD-ATP-binding cassette (PD-ABC) protein form #2.
DE PD-ATP-binding cassette; PD-ABC; chromosome 19p13.3; spleen; thymus;
KW peripheral blood leukocyte; bone marrow; lymph node; dyslipidaemia;
KW cardiovascular disorder; inflammatory disorder; abnormal calcium flux;
KW epilepsy; coronary artery disease; tangier's disease; atherosclerosis;
KW familial high-density lipoprotein deficiency; fatty liver disease;
KW atherosclerosis; diabetes; insulin resistance; obesity; drug screening;
KW alcoholism; retinal degeneration; hypertension; vascular disease.
XX XX
OS Homo sapiens.
XX XX
XX WO200153490-A1.
XX XX
XX 26-JUL-2001.
XX XX
XX 23-JAN-2001; 2001WO-US002191.
XX XX
XX 24-JAN-2000; 2000US-0177889P.
XX 30-JUN-2000; 2000US-0215405P.
XX XX
XX (WARN) WARNER LAMBERT CO.
XX XX
PI Johns MA, Tafuri SR, Wang M;
XX WPI; 2001-442259/47.
XX DR N-PSDB; AAS08707.
XX XX
XX New Human PD-ABC DNA molecules and proteins for diagnosis and treatment
XX of dyslipidemia, epilepsy and diseases related to abnormal calcium flux.
XX Claim 10; Page 64-72; 77pp; English.
XX XX
XX The sequence represents human PD-ATP-binding cassette (PD-ABC) protein
XX form 2. PD-ABC maps to chromosome 19p13.3 and is expressed in various
XX tissues including spleen, thymus, peripheral blood leukocytes, bone
XX marrow and lymph nodes. The PD-ABC DNA molecules and proteins are used to
XX diagnose and treat cardiovascular disorders, inflammatory disorders,
XX dyslipidaemia, epilepsy, diseases related to abnormal calcium flux,
XX coronary artery disease, tangier's disease, familial high-density
XX lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
XX insulin resistance, obesity, alcoholism, retinal degeneration,
XX hypertension and vascular disease. The sequences are also used in drug
XX screening assays
XX Sequence 1873 AA;
SQ

Query Match 84.9%; Score 9460.5; DB 4; Length 1873;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1835; Conservative 6; Mismatches 27; Indels 15; Gaps 4;
Qy 1 MAFWTQMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHHPLEBHECHFFNKP 60
Db 1 MAFWTQMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHHPLEBHECHFFNKP 60
Qy 61 LPSAGTVPMVLOGLICNNVNTCFPOLTPGEBPGRLSNFNDLSVSRLLADARTVLGGASAH 120
Db 61 LPSAGTVPMVLOGLICNNVNTCFPOLTPGEBPGRLSNFNDLSVSRLLADARTVLGGASAH 120
Qy 121 TLAGLGKLIATLAARSTAOPTKQSPLEPPMDVAELTSLRTESIGLALGQAQEP 180
Db 121 TLAGLGKLIATLAARSTAOPTKQSPLEPPMDVAELTSLRTESIGLALGQAQEP 180
Qy 181 HSLEAAEDLAQELLALRSLVELRALQRPRTSGPLELSEALCSVRGFSSTVGP 240
Db 181 HSLEAAEDLAQELLALRSLVELRALQRPRTSGPLELSEALCSVRGFSSTVGP 240
Qy 241 YEASDLMELVGPESALPDSSISLPACSELIGALDSHPILSRLLWRRUKPLILGKLLFAP 300
Db 241 YEASDLMELVGPESALPDSSISLPACSELIGALDSHPILSRLLWRRUKPLILGKLLFAP 300
Qy 301 TPTRKLMACVNRTFEELTLRLDVREVMELGPRIFTFMNDSSNVAMLQRLQMDGRR 360
Db 301 TPTRKLMACVNRTFEELTLRLDVREVMELGPRIFTFMNDSSNVAMLQRLQMDGRR 360
Qy 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAPSEA 420
Db 361 QPRPGGRDHMEALRSFLDPGSGGYSWQDAHADVGHVLTGLRVTECLSLDKLEAPSEA 420
Qy 421 LVSRALQLLAHFRWAGVFLPDESDSDPTEHTPDLPDGHVRKIRMDIDVVRTNKIR 480
Db 421 LVSRALQLLAHFRWAGVFLPDESDSDPTEHTPDLPDGHVRKIRMDIDVVRTNKIR 480
Qy 481 DRFWDGPAADPLTDLRYVMGGFYLLQDLVERAAVRVLSGANPRAGLYLQOMPVCYVDD 540
Db 481 DRFWDGPAADPLTDLRYVMGGFYLLQDLVERAAVRVLSGANPRAGLYLQOMPVCYVDD 540
Qy 541 VFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKRETLRDLTMRAVGLSRAVLWLGWFLSCL 600
Db 541 VFLVLSRSLPLFLTLAWIYSVTLTVKAVVREKRETLRDLTMRAVGLSRAVLWLGWFLSCL 600
Qy 601 GPFLLSAAALLVVLKGDILPYSHPGVFLFLAFAVATVTSQFLLSAFSTRANLAAACG 660
Db 601 GPFLLSAAALLVVLKGDILPYSHPGVFLFLAFAVATVTSQFLLSAFSTRANLAAACG 660
Qy 661 GLAYFSLYPLVLCVARDRLPAGGRVAASLLSPVAFGFCESLALLEOEGGAQWNVG 720
Db 661 GLAYFSLYPLVLCVARDRLPAGGRVAASLLSPVAFGFCESLALLEOEGGAQWNVG 720
Qy 721 TRPTADVFLSAQVSGLLLDLDAALYGLATWYLCVPGQYGIPEPWNFFRRSYWCGPRPP 780
Db 721 TRPTADVFLSAQVSGLLLDLDAALYGLATWYLCVPGQYGIPEPWNFFRRSYWCGPRPP 780
Qy 781 KSPAPCPTDPDKVLEVEAPPGLSPGVSRSLERKPFPGSPQALRGSLDLPYQGHITAF 840
Db 781 KSPAPCPTDPDKVLEVEAPPGLSPGVSRSLERKPFPGSPQALRGSLDLPYQGHITAF 840
Qy 841 GHNGAGKTTTSLISGLFPSPGGSATILGHDVRSMAAIRPHLGVCQYNNVLPMLTVDE 900
Db 841 GHNGAGKTTTSLISGLFPSPGGSATILGHDVRSMAAIRPHLGVCQYNNVLPMLTVDE 900
Qy 901 HVWFYGLKGLSAAVVGPEODRLIQDVLGVSQSVQTRHLSGGMQRKLSVAIAFVGSGQV 960
Db 901 HVWFYGLKGLSAAVVGPEODRLIQDVLGVSQSVQTRHLSGGMQRKLSVAIAFVGSGQV 960
Qy 961 VILDEPTAGVDPAARRGIWELLKLYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Db 961 VILDEPTAGVDPAARRGIWELLKLYREGRTLILSTHLLDEALLGDRVAVVAGRLCCCG 1020
Qy 1021 SPLFLRRHLGGVYTLVVKARLPLTTNEKADTMEGSDVTQREKNGSQGSRVGTQPOLLA 1080

Db 1021 SPLFLRLHSGSYLTAVKARLPLTTNEKADTDMEGSVDTQKXNGSQSVGTQPLIA 1080
QY 1081 LVQHWVPGARLVEELPHELVVLVLYTGAHDGSGFATLFRDLRLAELRLTGVGSDTSLE 1140
Db 1081 LVQHWVPGARLVEELPHELVVLVLYTGAHDGSGFATLFRDLRLAELRLTGVGSDTSLE 1140
QY 1141 EIFLKVBECAADTDMEDGSCQHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
Db 1141 EIFLKVBECAADTDMEDGSCQHLCTGTAGLDVTLRLKMPQETALENGEPAGSAPETD 1200
QY 1201 QSGGPDVGRVQGWALTRQOLQALLKRPFLARRSRRGLFAQIVLPALFVGLALVFSILV 1260
Db 1201 QSGGPDVGRVQGWALTRQOLQALLKRPFLARRSRRGLFAQIVLPALFVGLALVFSILV 1260
QY 1261 PPFHYPALRLSPTMYGAQVSPFSDAPGDCGRARLEALLQEAELGEPVVOHSHRFS 1320
Db 1261 PPFHYPALRLSPTMYGAQVSPFSDAPGDCGRARLEALLQEAELGEPVVOHSHRFS 1320
QY 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVEVVQ 1380
Db 1321 PEVPAEAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGPPPPQAVTGSVEVVQ 1380
QY 1381 NLTGRLSDFLVKTYPRLVROGLTKKWNVRYGSGFSLGGRDPGLPSQELGRSVEELW 1440
Db 1381 NLTGRLSDFLVKTYPRLVROGLTKKWNVRYGSGFSLGGRDPGLPSQELGRSVEELW 1440
QY 1441 ALLSPLPGGALDRVLKNTTAWAHSIDAQSLKIWFNNKGWSWAFVNRASNAIIRAHLP 1500
Db 1441 ALLSPLPGGALDRVLKNTTAWAHSIDAQSLKIWFNNKGWSWAFVNRASNAIIRAHLP 1500
QY 1501 PGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLSICVVPFAMSVFPAFTLVLEEE 1560
Db 1501 PGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLSICVVPFAMSVFPAFTLVLEEE 1560
QY 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNVYLVPACIVVLIFLAFQRAYVAPANLPALL 1620
Db 1561 RVTRAKHLQMGSLPTLYWLNFLWDMCNVYLVPACIVVLIFLAFQRAYVAPANLPALL 1620
QY 1621 LLLLYGWSITPLMYPASFFSVPTAYVLTCLNLTGINGSMTATVLELFSQKLOEV 1680
Db 1621 LLLLYGWSITPLMYPASFFSVPTAYVLTCLNLTGINGSMTATVLELFSQKLOEV 1680
QY 1681 SRLKQVFLIPHFCLGRGLDMVNRQAMADAFERLGRDQFOSPLRVEWVGKLLAMVIQ 1740
Db 1681 SRLKQVFLIPHFCLGRGLDMVNRQAMADAFERLGRDQFOSPLRVEWVGKLLAMVIQ 1740
QY 1741 GPLFLFTLLQHRSQLLPQPRVRSPLPLGEEDEVDARERVRVQATGQDVLVRLNLT 1800
Db 1741 GPLFLFTLLQHRSQLLPQPRVRSPLPLGEEDEVDARERVRVQATGQDVLVRLNLT 1800
QY 1801 VYRGORMPAVDRLCLGIPGECFGLLVNGAGKTSFRMVTGDTLASRGEAVLAGHSVAR 1860
Db 1801 VYRGORMPAVDRLCLGIPGECFGLLVNGAGKTSFRMVTGDTLASRGEAVLAGHSVAR 1860
QY 1861 EPSAAHLGMYCPOSDAIFELL 1882
Db 1855 -----HLLLYTCM---PIFVLL 1869

RESULT 12

ABG72695

ID ABG72695 standard; protein; 2167 AA.

XX AC ABG72695;

XX AC ABG72695;

XX 10-MAR-2003 (first entry)

XX 10-MAR-2003 (first entry)

DE Mouse ATP-binding cassette transporter-like protein, ABCL.

XX Mouse; ATP-binding cassette transporter-like protein; ABCL;

KW lipid transport; cardiovascular disease; hypertriglyceridaemia;

KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;

nervous system disorder; Stargardt disease; degenerative disorder;
inflammatory retinopathy; cystic fibrosis; multidrug resistance;
lymphoid condition; myeloid cell condition; AIDS; lymphoma;
acquired immunodeficiency disorder; leukaemia; neutropenia; anaemia;
autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
hypothalamus disorder; obesity; diabetes; reproductive disorder;
energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
autoimmune disease; inflammatory disease; multiple sclerosis.

Mus musculus.

Key Location/Qualifiers
Peptide 1..46
Protein /label= Signal_peptide
47..2167
/label= Mature_ABL

US2002127647-A1.

XX 12-SEP-2002.

XX 28-NOV-2001; 2001US-00995542.

XX 28-NOV-2000; 2000US-0253520P.

XX (SHUT/) SHUTTER J.

XX (ULIA/) ULIAS L.

XX Shutter J, Ulias L;

XX WPI; 2003-147394/14.

XX N-PSDB; ABX14665.

Novel ATP-binding cassette transporter-like polypeptides and
polynucleotides useful for diagnosing, preventing, treating disorders
involving immune, nervous system, thyroid, hypothalamus and impaired
transport of lipids.

Claim 13; Fig 1; 149pp; English.

The invention relates to an isolated murine and human ATP-binding
cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
3111. Also include are the nucleic acids encoding the ABCL proteins,
vectors, host cells, ABCL binding agents, a selective binding agent or
its fragment comprising at least one complementarity determining region
(CDR) with specificity for ABCL which (produced by immunising an animal
with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
fusion polypeptide, a device comprising a membrane suitable for
implantation (permeable to the protein and impermeable to materials
detrimental to the cells, and cells encapsulated within the membrane)
where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
and modulators are useful for the diagnosis and/or treatment of diseases
and conditions involving impaired transport of lipids, including
cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
involving functional and trophic disturbances of the nervous system such
as Stargardt disease, degenerative and inflammatory retinopathy, cystic
fibrosis, conditions involving multidrug resistance, conditions involving
lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,
neutropenia, anaemia and autoimmune diseases, conditions involving the
thyroid e.g. hyper and hypothyroidism; conditions involving the
hypothalamus including obesity, diabetes, reproductive disorders, energy
balance disorders, peripheral neuropathies including myelinopathies and
axonopathies, autoimmune and inflammatory diseases involving the nervous
system including multiple sclerosis. The present sequence represents
murine ABCL

Sequence 2167 AA;

Query Match 77.3%; Score 8613.5; DB 6; Length 2167;
Best Local Similarity 76.7%; Pred. No. 0;

QY 2118 VGVDPAQLOHPKRVYSQFLDDPSTAEITVL 2146
DB 2139 BEEVSKRGQHPKRVSRFLEDPSSVETMI 2167
RESULT 13
ID ABG72697
AC ABG72697;
XX 10-MAR-2003 (first entry)
XX Human ATP-binding cassette transporter-like protein, ABCI1550.
XX Human; ATP-binding cassette transporter-like protein; ABCI, ABCI1550;
KW lipid transport; cardiovascular disease; hypertriglyceridaemia;
KW atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;
KW nervous system disorder; Stargardt disease; degenerative disorder;
KW inflammatory retinopathy; cystic fibrosis; multidrug resistance;
KW lymphoid condition; myeloid cell condition; AIDS; lymphoma;
KW acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;
KW autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;
KW hypothalamus disorder; obesity; diabetes; reproductive disorder;
KW energy balance disorder; peripheral neuropathy; myelinopathy; axonopathy;
KW autoimmune disease; inflammatory disease; multiple sclerosis.
XX Homo sapiens.
XX OS
XX PN US2002127647-A1.
XX XX
XX 12-SEP-2002.
XX PF 28-NOV-2001; 2001US-00995542.
XX PR 28-NOV-2000; 2000US-0253520P.
XX PA (SHUT/) SHUTTER J.
XX PI (ULIA/) ULIAS L.
XX PI Shutter J, Ulias L;
XX WPI; 2003-147394/14.
DR N-PSDB; ABX14667.
XX XX
XX Novel ATP-binding cassette transporter-like polypeptides and
PT polynucleotides useful for diagnosing, preventing, treating disorders
PT involving immune, nervous system, thyroid, hypothalamus and impaired
PT transport of lipids.
XX PS
PS Claim 13; Fig 3; 149pp; English.
XX XX
XX The invention relates to an isolated murine and human ATP-binding
CC cassette transporter-like (ABCL) polypeptide, or the amino acid sequence
CC encoded by the DNA insert in ATCC Deposit Nos PTA-3109, PTA-3110 or PTA-
CC 3111. Also include are the nucleic acids encoding the ABCL proteins,
CC vectors, host cells, ABCL binding agents, a selective binding agent or
CC its fragment comprising at least one complementarity determining region
CC (CDR) with specificity for ABCL which (produced by immunising an animal
CC with ABCL), a hybridoma that produces the CDR, viral vectors, an ABCL
CC fusion polypeptide, a device comprising a membrane suitable for
CC implantation (permeable to the protein and impermeable to materials
CC detrimental to the cells, and cells encapsulated within the membrane)
CC where the cells secrete ABCL, an ABCL transgenic non-human mammal and an
CC array of ABCL nucleic acid molecules. The ABCL polypeptide, nucleic acids
CC and modulators are useful for the diagnosis and/or treatment of diseases
CC and conditions involving impaired transport of lipids, including
CC cardiovascular disease, hypertriglyceridaemia, atherosclerosis,
CC hypercholesterolaemia, Tangier disease, dyslipidaemias, conditions
CC involving functional and trophic disturbances of the nervous system such
CC as Stargardt disease, degenerative and inflammatory retinopathy, cystic
CC fibrosis, conditions involving multidrug resistance, conditions involving
CC lymphoid and myeloid cells, including AIDS, lymphomas, leukaemias,

CC neutropaenia, anaemia and autoimmune diseases, conditions involving the
CC thyroid e.g. hyper and hypothyroidism; conditions involving the
CC hypothalamus including obesity, diabetes, reproductive disorders, energy
CC balance disorders, peripheral neuropathies including myelinopathies and
CC axonopathies, autoimmune and inflammatory diseases involving the nervous
CC system including multiple sclerosis. The present sequence represents
CC human ABCL truncated variant, ABCI1550
XX
SQ Sequence 1550 AA;
Query Match 68.2%; Score 7596; DB 6; Length 1550;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1461; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 167 ESLGALGQAQEPHLSLLEAEADLAQELLALSVELRALLORPRTSGPSELLSALCS 226
DB 29 ESLGALGQAQEPHLSLLEAEADLAQELLALSVELRALLORPRTSGPSELLSALCS 88
QY 227 VRGPSTVGPSSLNWEASDLMELVGOEPESALPDSLSLSPACSELIGALDSHPLSLRLLWR 286
DB 89 VRGPSTVGPSSLNWEASDLMELVGOEPESALPDSLSLSPACSELIGALDSHPLSLRLLWR 148
QY 287 LKPLILGKLLFAPDPTFTRKLMQVNRTEELTLRDVREVMELGPRIFTFMDSSNVA 346
DB 149 LKPLILGKLLFAPDPTFTRKLMQVNRTEELTLRDVREVMELGPRIFTFMDSSNVA 208
QY 347 MLQRLQMDGRRORPGRGDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTRC 406
DB 209 MLQRLQMDGRRORPGRGDHMEALRSFLDPGSGYSWQDAHADVGHVLTGRVTRC 268
QY 407 LSLDKLEAAPSEAAALVSRALQLAEHREFWAGVYFLGPDSSDPTBHTPTDLGFGHVRKI 466
DB 269 LSLDKLEAAPSEAAALVSRALQLAEHREFWAGVYFLGPDSSDPTBHTPTDLGFGHVRKI 328
QY 467 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWCGFYVLQDLVERAAVRLSGANPRAG 526
DB 329 RMDIDVTRTKIRDRFWDGPAADPLTDLRYVWCGFYVLQDLVERAAVRLSGANPRAG 388
QY 527 LYLOQMPYPCYVDDVFLRVLSRSLPLTLTAWIYSVTLTVKAVREKETRLDTRWAMGL 586
DB 389 LYLOQMPYPCYVDDVFLRVLSRSLPLTLTAWIYSVTLTVKAVREKETRLDTRWAMGL 448
QY 587 SRAVLWLGWFLSCIGPFLLSAALLVLKLGDLIPYSHPGVYFLFAAFATVTSFLL 646
DB 449 SRAVLWLGWFLSCIGPFLLSAALLVLKLGDLIPYSHPGVYFLFAAFATVTSFLL 508
QY 647 SAFTSRANLAAACGLAYFSLYLYPVLCVAVNRDRLPAGGRVAASLLSPVAFGCGESLAL 706
DB 509 SAFTSRANLAAACGLAYFSLYLYPVLCVAVNRDRLPAGGRVAASLLSPVAFGCGESLAL 568
QY 707 LEEQGEQAHNVGTRPTADVPFLAQVSGLLLLDAAALYGLATWYLEAVCPGQYGIPEPWN 766
DB 569 LEEQGEQAHNVGTRPTADVPFLAQVSGLLLLDAAALYGLATWYLEAVCPGQYGIPEPWN 628
QY 767 FPFRRSYWCGPRPKSPACPTPLDPKVLVEAEPEGLSPGVSVRSLEKRFPGSPQALRG 826
DB 629 FPFRRSYWCGPRPKSPACPTPLDPKVLVEAEPEGLSPGVSVRSLEKRFPGSPQALRG 688
QY 827 LSLDFYQGHITAFLHNGAGKTTTILSLGSLFPSPGSGSAFILGHVDVRSMAAIRPHLGVC 886
DB 689 LSLDFYQGHITAFLHNGAGKTTTILSLGSLFPSPGSGSAFILGHVDVRSMAAIRPHLGVC 748
QY 887 PQYNVLFDMLTVDHFWVYGRKLGISAUVGPEODRLLODVLVSKQSVQTRHLSGGWQR 946
DB 749 PQYNVLFDMLTVDHFWVYGRKLGISAUVGPEODRLLODVLVSKQSVQTRHLSGGWQR 808
QY 947 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIIWELLKYREGRTILILSTHLDLDEALLGD 1006
DB 809 KLSVAIAFVGSGQVVLDEPTAGVDPASRRGIIWELLKYREGRTILILSTHLDLDEALLGD 868
QY 1007 RVAVVAGRLCCCGSPFLRRHLSGGYLYTLVKARLPLTTNEKADTMEGSVDTROEKN 1066
DB 869 RVAVVAGRLCCCGSPFLRRHLSGGYLYTLVKARLPLTTNEKADTMEGSVDTROEKN 928


```
QY 1067 GSQSRVGTGTPOLLALVOHWVPGARLVEELPHELVLVLPYTGAGHDSFATLFRDLTDLAE 1126
D 929 GSQSRVGTGTPOLLALVOHWVPGARLVEELPHELVLVLPYTGAGHDSFATLFRDLTDLAE 988
QY 1127 LRLTGYGISTSLBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPPOETA 1186
D 989 LRLTGYGISTSLBEIFLKVVEECAADTDMEDGSCGQHLCTGIAGLDVTLRLKMPPOETA 1048
QY 1187 LENGEAGSAPETDQSGPDAGVRQGWALTRQOLQALLKRFLLARRRRGLFAQIVLP 1246
D 1049 LENGEAGSAPETDQSGPDAGVRQGWALTRQOLQALLKRFLLARRRRGLFAQIVLP 1108
QY 1247 ALFVGLALVLSLVPFGHYPALRLSPMYGAOVFFSEADPGDGRARLLEALLQAGL 1306
D 1109 ALFVGLALVLSLVPFGHYPALRLSPMYGAOVFFSEADPGDGRARLLEALLQAGL 1168
QY 1307 EEPVQVHSHRFSAPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1366
D 1169 EEPVQVHSHRFSAPAEVAKVLASGNWTPESPSPACQSQPGARRLLPDCPAAAGGP 1228
QY 1367 PPOAVTGSSEVQNLGRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSIGGRDPGL 1426
D 1229 PPOAVTGSSEVQNLGRNLSDFLVKTYPRLVQGLTKKWNVEVRYGFSIGGRDPGL 1288
QY 1427 PSGQELGRSVEELWALLSPPLGGALDRVLKNTLTAHSLDAQDSLKTWNNKGHWSVAF 1486
D 1289 PSGQELGRSVEELWALLSPPLGGALDRVLKNTLTAHSLDAQDSLKTWNNKGHWSVAF 1348
QY 1487 VNRASNAIRLAHLPFGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLSVCVVFAM 1546
D 1349 VNRASNAIRLAHLPFGPARHAHSITTLNHLNLTKEQLSEALMASSVDVLSVCVVFAM 1408
QY 1547 SFVPASFTVLIERVTRAKHLQMGSLPFLYWLGNFLMDMCMNYLVPACIVLIFLAFQ 1606
D 1409 SFVPASFTVLIERVTRAKHLQMGSLPFLYWLGNFLMDMCMNYLVPACIVLIFLAFQ 1468
QY 1607 QRAVAPANLPAMLLLLLVGWSITPL 1633
D 1469 QRAVAPANLPAMLLLLLVGWSITPL 1495
RESULT 14
AAB38111
ID AAB38111 standard; protein; 2261 AA.
AC AAB38111;
XX
XX
XX
XX
XX 29-JAN-2001 (first entry)
XX
XX Human ABC1 cholesterol transporter mutant, V771M.
XX
XX Human ABC1 cholesterol transporter; chromosome 9q31;
XX ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
XX Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
XX cardiovascular disease; coronary artery disease; coronary restenosis;
XX cerebrovascular disease; peripheral vascular disease;
XX Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
XX X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
XX progonis; prophylaxis; drug screening; transgenic animal; mutant;
XX mutin.
XX
XX Homo sapiens.
XX
XX WO200055318-A2.
XX
XX 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-IB000532.
XX
XX 15-MAR-1999; 99US-0124702P.
XX 08-JUN-1999; 99US-0138048P.
XX 17-JUN-1999; 99US-0139600P.
```

```
PR 01-SEP-1999; 99US-0151977P.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
XX Hayden MR, Wilson AR, Fimstone SN;
XX WPI; 2000-587528/55.
XX
XX New ABC1 polypeptide is useful for treating diseases associated with ABC1
XX biological activity, e.g. Alzheimer's disease, Huntington's disease and
XX cancer.
XX
XX Example; Page; 229pp; English.
XX
XX The invention relates to the human ABC1 cholesterol transporter protein
XX (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
XX a member of the ATP-binding cassette (ABC transporter) superfamily of
XX proteins, and plays a crucial role in cholesterol transport, particularly
XX intracellular cholesterol trafficking in monocytes and fibroblasts, being
XX involved in cholesterol efflux from the cell. The gene encoding ABC1 is
XX located on chromosome 9q31, and mutations in this gene are associated
XX with two genetic HDL (high density lipoprotein) deficiency disorders,
XX Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
XX are distinguishable in that TD is an autosomal recessive disorder, while
XX FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
XX cholesterol") in the blood correlate with a high risk of cardiovascular
XX disease, particularly coronary artery disease, but also cerebrovascular
XX disease, coronary restenosis, and peripheral vascular disease.
XX Conversely, a high level of HDL has protective effects against
XX cardiovascular disease. The invention provides genetic constructs and
XX transgenic cells and non-human animals comprising human ABC1 nucleic
XX acids, and methods of gene therapy for the treatment or prevention of
XX cardiovascular disease comprising the administration of an expression
XX vector encoding ABC1 or an active fragment thereof. The invention also
XX encompasses compounds which mimic ABC1 activity, compounds which
XX stimulate ABC1 expression and methods of screening for such compounds. It
XX further relates to methods for determining whether a patient has an
XX increased risk for cardiovascular disease due to polymorphisms in the
XX ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or
XX prevent cardiovascular disease, especially coronary artery disease,
XX cerebrovascular disease, coronary restenosis or peripheral vascular
XX disease. They may also be used in the treatment of diseases associated
XX with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
XX disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
XX The invention specifically excludes proteins with the exact amino acid
XX sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
XX acid with the exact sequence as GenBank Accession No: A7012376.1. The
XX present sequence represents a mutant human ABC1 cholesterol transporter
XX associated with an altered cholesterol level and therefore an altered
XX risk of cardiovascular disease. Note: The present sequence is not shown
XX in the specification, but is derived from the native human ABC1 shown on
XX pages 152-157
XX
XX Sequence 2261 AA;
```

```
Query Match 51.8%; Score 5769.5; DB 3; Length 2261;
Best Local Similarity 50.3%; Pred. No. 0;
Matches 1147; Conservative 368; Mismatches 606; Indels 161; Gaps 26;

QY 1 MAFWTQLMLLLWKNFMVRRQPVQLLVLLWLPFLFFILVAVRHSHPLEHCHFFNKP 60
D 1 MACWPQRLLLWKNLTFRRRQTCLLLEVAWPLFILLISVRLSYFPYQHECHFFNKA 60
QY 61 LPSAGTVPWLQGLICNVNNTCFPOLTEGEPGRLSNFNDSLVSRLLDARTVLGGASAH 120
D 61 MPSAGTLPWVQGIICNANNPCFRYPTGEAPGVGVNFNKSIVARLFSRARLLYSQKDT 120
QY 121 TLAGEGLKIATLRAARSTAQ-----PQTKQSPLEPPML-- 154
D 121 SMKDMRKVJLTQQIKKSSNLLKLQDFLVNDETSGFLYHNLSPKSTVDKMLRADVILH 180
QY 155 -----DVAEL-----LTSLL 164
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:57:46 ; Search time 27508 Seconds
(without alignments)
11696.929 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctccggggcgcgcgtccccc.....cgtgctggtgaaaaa 6804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6804	100.0	6804	6	AX644622	AX644622 Sequence
2	6762.4	99.4	6791	6	AX686999	AX686999 Sequence
3	6717.8	98.7	6768	6	AX320362	AX320362 Sequence
4	6710	98.6	6795	6	AX429481	AX429481 Sequence
5	6686.4	98.3	6704	9	AF328787	AF328787 Homo sapi
6	6565.6	96.5	6588	9	AF250238	AF250238 Homo sapi
7	6478.4	95.2	6522	6	AX202218	AX202218 Sequence
8	6398.4	94.0	6432	6	AX320364	AX320364 Sequence
9	5928.2	87.1	6027	9	AB055390	AB055390 Homo sapi
10	5557.2	81.7	6324	6	CQ719367	CQ719367 Sequence
11	5458.2	80.2	5669	6	AX202220	AX202220 Sequence
12	4381.6	64.4	4653	6	AX644625	AX644625 Sequence
13	3902.6	57.4	6513	10	AB097814	AB097814 Rattus no
14	3902.4	57.4	6633	6	AX644619	AX644619 Sequence
15	3885.8	57.1	6607	6	AX080493	AX080493 Sequence
16	3885.8	57.1	6607	10	AF287141	AF287141 Mus muscu
17	3303	48.5	5762	6	AX080462	AX080462 Sequence
18	2861.8	42.1	2911	6	BD234208	BD234208 ATP-bind
19	1808.6	26.6	1879	6	AX834707	AX834707 Sequence

20	1808.6	26.6	1879	9	AK097344	AK097344 Homo sapi
21	1788.4	26.3	7074	5	AF362377	AF362377 Gallus ga
22	1740.6	25.6	6786	9	AB055982	AB055982 Homo sapi
23	1739	25.6	10412	6	CQ785966	CQ785966 Sequence
24	1737.4	25.5	7860	6	AR393480	AR393480 Sequence
25	1737.4	25.5	7860	6	AX092594	AX092594 Sequence
26	1737.4	25.5	10442	6	AX060713	AX060713 Sequence
27	1737.4	25.5	10442	6	AX060892	AX060892 Sequence
28	1735.8	25.5	10474	9	AF285167	AF285167 Homo sapi
29	1735.8	25.5	10474	6	AX060719	AX060719 Sequence
30	1735.8	25.5	10474	6	AX060721	AX060721 Sequence
31	1735.8	25.5	10474	6	AX060898	AX060898 Sequence
32	1735.8	25.5	10474	6	AX060900	AX060900 Sequence
33	1734.2	25.5	6880	6	BD234191	BD234191 ATP-bind
34	1734.2	25.5	6880	6	AX253450	AX253450 Sequence
35	1734.2	25.5	6880	9	HS012376	HS012376 Homo sapi
36	1734.2	25.5	7260	6	AX253452	AX253452 Sequence
37	1734.2	25.5	9497	6	AX059973	AX059973 Sequence
38	1734.2	25.5	9497	9	AF165281	AF165281 Homo sapi
39	1734.2	25.5	9741	6	AX127830	AX127830 Sequence
40	1734.2	25.5	9741	6	AX139817	AX139817 Sequence
41	1734.2	25.5	9741	6	AX351038	AX351038 Sequence
42	1734.2	25.5	9854	6	AX127831	AX127831 Sequence
43	1734.2	25.5	9854	6	AX139818	AX139818 Sequence
44	1725.4	25.4	7878	10	MMABC1	X75926 Mus musculu
45	1724.4	25.3	6801	10	AY208182	AY208182 Rattus no

ALIGNMENTS

RESULT 1
AX644622
LOCUS AX644622 6804 bp DNA linear PAT 27-FEB-2003
DEFINITION Sequence 4 from Patent WO02099108.
ACCESSION AX644622
VERSION AX644622.1 GI:28610637
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shutter, J. and Ulias, L.
TITLE Atp-binding cassette transporter-like molecules and uses thereof
JOURNAL Patent: WO 02099108-A 4 12-DEC-2002;
Angen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..6804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
210..6650
/note="unnamed protein product"

1..6804
/codon_start=1
/protein_id="CAD68132.1"
/db_xref="GI:28610638"
/translation="MAFWQLMLLNKFNRRQPVQLLVLLWLPFLFLLVAVRH
SHPLEHECHFPNKLPLPSAGTVPLQLICNVNNTCFPQLTPGEEFGRLSNFNDLSV
SRLADARTVLGGASAHRTLAGKLIATLRAAFSTAQPOKQSPLEPMMLDVAELL
TSLLRTSLGLAQAOEPLHLEAAEDLAQELLALRLVLRALLQRPGRGTGPLE
LLSEALCSVRGSPSTVGSINWYEASDLMELVGPESALPDSLSLSPACSLIGALDS
HIFSLRLRLKLLILGLKLPADPTPTKLMQAVNRTFELTLRLRYREWEMLGPR
IFTFMDSNVNMLQRLQMDGRRQPRFGRRHMEALRSLDFGSGGYSWQDAHAD
VHGLVGTGRTVECLSLDKLEAPSEAAALSRALQDLAHRFWAGVFLVGEVSSDPT
EGHTPDLPGLGHVRIKIRMDIDVTRTKIRDRFMDPGFAADPLTDLRYVMGFGVLLQD
LVRAAVRSLRGNPRAGLYLQOMPYPYVDVFLVLSRLPLFLTLAWYVTLTV
KAVUREKETRLRDTMRAWGLSRVLMGLWFLSCIGPLLSAALLVLVLKGLDILPYSH
PGVVFPLAPAVATVYQSLLSAFFSRANLAACGGLAFSLYLPVYLCVAMWRDLRP
AGRVAAASLLSPVAFGCGESLALLEEGGGAQHNWVGTPTADVFSLAQVSGULLLD
AALYGLATWYLEAVCPQYGIPIPPNFPFRSPKPPKSPACPTDLQPKVLEE
APGLSPGVSRVSLRKPFGSPQALRGLSLDFYQGHITAFGLHNGAGKTTLLSG
LFPFGSGSAFILGHVRSMAAIRPLHGVCPQYNVLFDMLTVDHEHWFYGRKLGSAA

Db 6121 TGTGCGGAGGCGCGTTCAGTGATGCTCACCTCCATAGCATGAGGAGTGTGAAGCGC 6180
Qy 6181 TCTGCTCCGCGCTAGCCATCATGTGTAATGGCGCGTTCGGTGCCTGGCGAGCCGCAAC 6240
Db 6181 TCTGCTCCGCGCTAGCCATCATGTGTAATGGCGCGTTCGGTGCCTGGCGAGCCGCAAC 6240
Qy 6241 ATCTCAAGGGCAGATTCGCGGCGGTCTACACACTGACCTCGGGTGCCTCCGCAAGGT 6300
Db 6241 ATCTCAAGGGCAGATTCGCGGCGGTCTACACACTGACCTCGGGTGCCTCCGCAAGGT 6300
Qy 6301 CCCAGCGGCGAGCGGCTTCGTGGCGGCGAGTTCCTTGGTTCGAGTTCGCGAGGCGAC 6360
Db 6301 CCCAGCGGCGAGCGGCTTCGTGGCGGCGAGTTCCTTGGTTCGAGTTCGCGAGGCGAC 6360
Qy 6361 ATGAGGCGCGCTGCGCTTCAGCTGCGCGGAGGCGCTGCGCCCTGCGCGCGCTCT 6420
Db 6361 ATGAGGCGCGCTGCGCTTCAGCTGCGCGGAGGCGCTGCGCCCTGCGCGCGCTCT 6420
Qy 6421 TTGAGAGCTGCGGTGCAACGCGGCGAGCAGCGCGTGGAGACTTTTCCGTGAGCCAGA 6480
Db 6421 TTGAGAGCTGCGGTGCAACGCGGCGAGCAGCGCGTGGAGACTTTTCCGTGAGCCAGA 6480
Qy 6481 CGATGCTGGAGAGGTATTTCTGTACTTCTCCAGGACGAGGAGGAGGACGACCG 6540
Db 6481 CGATGCTGGAGAGGTATTTCTGTACTTCTCCAGGACGAGGAGGAGGACGACCG 6540
Qy 6541 AAGAGCAGAAAGGAGGAGGTGGAGTGGAGTGGACCCCGCGGAGGCTGCGAGCAGCCCAAC 6600
Db 6541 AAGAGCAGAAAGGAGGAGGTGGAGTGGAGTGGACCCCGCGGAGGCTGCGAGCAGCCCAAC 6600
Qy 6601 CGCTCAGCAGTCTTCGTATGACCTTAGCAGTGGCGAGACTGTCTCTGAGCCCTCCCTCC 6660
Db 6601 CGCTCAGCAGTCTTCGTATGACCTTAGCAGTGGCGAGACTGTCTCTGAGCCCTCCCTCC 6660
Qy 6661 CCTGGGGGCGCGGGAGGCGCTGGGAAATGGCAAGGCAAGGTAGAGTGGCTAGGAGCC 6720
Db 6661 CCTGGGGGCGCGGGAGGCGCTGGGAAATGGCAAGGCAAGGTAGAGTGGCTAGGAGCC 6720
Qy 6721 CTGAGCTCAGGCTGCGAGAGGCGTGGTGGCTGGAGAAATTAAGAGAGGCTGGAGAG 6780
Db 6721 CTGAGCTCAGGCTGCGAGAGGCGTGGTGGCTGGAGAAATTAAGAGAGGCTGGAGAG 6780
Qy 6781 AAGCCGTGCTGGTGAAGAAAAAAA 6804
Db 6781 AAGCCGTGCTGGTGAAGAAAAAAA 6804

RESULT 2
AX686999 LOCUS 6791 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 59 from Patent WO0204520.
ACCESSION AX686999
VERSION AX686999.1 GI:29409520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Raumann, B.E., Thornton, M., Ding, L., Yue, H., Tang, Y.T., Harland, L.,
Burford, N., Greene, B.D., Sanjanwala, M., Baughn, M.R., Yao, M.,
Yang, J., Patterson, C., Gandhi, A., Hafalia, A., Tribouley, C.,
Walia, N., Au-Young, J., Walsh, R.T., Ramkumar, J., Lu, Y., Lu, D.A.,
Azimzai, Y., Lal, P., Elliott, V., Nguyen, D., Xu, Y., Seilhamer, J.,
Borowsky, M., Kahn, F.A., Kearney, L., Thangavelu, K., Das, D. and
Polickey, J.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0204520-A 59 17-JAN-2002;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..6791
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7475603CB1"
ORIGIN
Query Match 99.4%; Score 6762.4; DB 6; Length 6791;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6780; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
Qy 12 CGCGCTCCCTGCTGCTGCGGCGGAGGCGCAAGAGCTGCGGAGCCCTGGAA 71
Db 1 CGCGCTCCCTGCTGCTGCGGCGGAGGCGCAAGAGCTGCGGAGCCCTGGAA 60
Qy 72 GAGCTTCAGGAACCTGCGCTGTGGGATAAAGGAATGAGGTTCAGAAAGGCGGAG 131
Db 61 GAGCTTCAGGAACCTGCGCTGTGGGATAAAGGAATGAGGTTCAGAAAGGCGCA 119
Qy 132 TTGCGCGGAGCGGACCGCAGCTTTCAGCCGAGCCGTTGCTGCTGCTGCTGCTGCT 191
Db 120 TTGCGCGGAGCGGACCGCAGCTTTCAGCCGAGCCGTTGCTGCTGCTGCTGCTGCT 179
Qy 192 CCCCGCGCAGCTCACCATGCGCTTCTGGGACACAGCTGATGCTGCTGCTGGAAGAT 251
Db 180 CCCCTGCCAGCTCACCATGCGCTTCTGGGACACAGCTGATGCTGCTGGAAGAT 239
Qy 252 TTCAATGATCCCGGAGACAGCCGCTCCAGCTCCCTGCTGCTGCTGCTGCTGCTGCT 311
Db 240 TTCAATGATCCCGGAGACAGCCGCTCCAGCTCCCTGCTGCTGCTGCTGCTGCTGCT 299
Qy 312 CTCTTCTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
Db 300 CTCTTCTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
Qy 372 CACTTCCCAACAAAGCACTGCGCATCGCGGCGGACCGTGGCTGCTGCTGCTGCTGCTGCT 431
Db 360 CACTTCCCAACAAAGCACTGCGCATCGCGGCGGACCGTGGCTGCTGCTGCTGCTGCTGCT 419
Qy 432 TGTAAATGTGAACAAACACTGCTTTCGCGAGCTGACACCGGCGGAGGAGCCCGGCGCTG 491
Db 420 TGTAAATGTGAACAAACACTGCTTTCGCGAGCTGACACCGGCGGAGGAGCCCGGCGCTG 479
Qy 492 AGCAACTTCAACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
Db 480 AGCAACTTCAACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Qy 552 GGGGCGAGTGGCCACAGGAGCTGCGCTGCGGCTGAGGAGCTGATCGCCACGCTGAGGCT 611
Db 540 GGGGCGAGTGGCCACAGGAGCTGCGCTGCGGCTGAGGAGCTGATCGCCACGCTGAGGCT 599
Qy 612 GCACGCGAGCGGCGGCGGAGCTTCAACCAACAGCAGCTTCCACCTGGAAACCAACCACTGCTG 671
Db 600 GCACGCGAGCGGCGGCGGAGCTTCAACCAACAGCAGCTTCCACCTGGAAACCAACCACTGCTG 659
Qy 672 GATGTCGGGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
Db 660 GATGTCGGGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Qy 732 CAAGCCAGGAGCGCTTGGCAGCTTCTGGAGGCGCTGAGGAGCTGAGGAGCTGAGGAGCTGCT 791
Db 720 CAAGCCAGGAGCGCTTGGCAGCTTCTGGAGGCGCTGAGGAGCTGAGGAGCTGAGGAGCTGCT 779
Qy 792 CTGCGCTGCGGAGCTTGGGAGCTTCTGGGAGCTTCTGGGAGCTTCTGGGAGCTTCTGGGAGCT 851
Db 780 CTGCGCTGCGGAGCTTGGGAGCTTCTGGGAGCTTCTGGGAGCTTCTGGGAGCTTCTGGGAGCT 839
Qy 852 GGCCCTCTGGAGTGTGTCAGAGGCGCTCTGCGAGTGTGAGGAGCTTCTGAGGAGCTTCTGAGGAGCT 911
Db 840 GGCCCTCTGGAGTGTGTCAGAGGCGCTCTGCGAGTGTGAGGAGCTTCTGAGGAGCTTCTGAGGAGCT 899
Qy 912 GGCCCTCTCTCAACTGCTAGGAGCTTCTGAGCTGAGGAGCTTCTGAGGAGCTTCTGAGGAGCTTCT 971
Db 900 GGCCCTCTCTCAACTGCTAGGAGCTTCTGAGCTGAGGAGCTTCTGAGGAGCTTCTGAGGAGCTTCT 959

972 GAATCCGCTTGCAGACAGCAGCTGAGCCCCCGCTGCTCGAGCTGATTGGAGCCCTG 1031
1032 GACAGCCA CCCGCTGTCCCCTGCTCTGAGACGCTGAAGCCTCTGATCTCTCGGGAAG 1091
1092 CTACTCTTTGACACAGATACACCTTTTACCCGGAAGCTATGCCCAAGGTGAACCGGACC 1151
1080 CTACTCTTTGACACAGATACACCTTTTACCCGGAAGCTATGCCCAAGGTGAACCGGACC 1139
1152 TTCGAGAGCTCACTCTGCTGAGGAGTATCCGCGAGGTGCGGAGATGCTGGGACCCCGG 1211
1140 TTCGAGAGCTCACTCTGCTGAGGAGTATCCGCGAGGTGCGGAGATGCTGGGACCCCGG 1199
1212 ATCTTCACTTTCATGAACGACAGTTCCTCAATGTGGCCATGCTGAGCGGCTCTGCGAGATG 1271
1200 ATCTTCACTTTCATGAACGACAGTTCCTCAATGTGGCCATGCTGAGCGGCTCTGCGAGATG 1259
1272 CAGGATGAAGGAAGAGCGACGCCAGACTGGAGGCGGGACACATGGAGGCCCTGCGA 1331
1260 CAGGATGAAGGAAGAGCGACGCCAGACTGGAGGCGGGACACATGGAGGCCCTGCGA 1319
1332 TCCTTTCTGGACCTCGGAGCGGTGCTACAGCTGGCAGAGCAGACACGCTGATGTGGG 1391
1320 TCCTTTCTGGACCTCGGAGCGGTGCTACAGCTGGCAGAGCAGACACGCTGATGTGGG 1379
1392 CACTTGGTGGGACGCTGGGCCGAGTGAAGGAGTGCCTGCTTGGACAAGCTGGAGGGG 1451
1380 CACTTGGTGGGACGCTGGGCCGAGTGAAGGAGTGCCTGCTTGGACAAGCTGGAGGGG 1439
1452 GCACCTTCAGAGCAGCGCTGTGTGCGGGCCCTGCAACTGCTGCGGGAACATCGATTTC 1511
1440 GCACCTTCAGAGCAGCGCTGTGTGCGGGCCCTGCAACTGCTGCGGGAACATCGATTTC 1499
1512 TGGGCGGCGCTGCTTCTTGGGACCTGAGGACTTTCAGACCCACAGAGCACCAAC 1571
1500 TGGGCGGCGCTGCTTCTTGGGACCTGAGGACTTTCAGACCCACAGAGCACCAAC 1559
1572 CCAGACTTGGGCCCGGCCACGTCGCGCATCAAAATCCGATGACATTTGACGTGTCAG 1631
1560 CCAGACTTGGGCCCGGCCACGTCGCGCATCAAAATCCGATGACATTTGACGTGTCAG 1619
1632 AGAACCAATGAATCAGGAGCAGGTTTGGGACCTTGGGACCTTGGGACCGCGGACCCCTGACC 1691
1620 AGAACCAATGAATCAGGAGCAGGTTTGGGACCTTGGGACCGCGGACCGCCCTGACC 1679
1692 GACCTGCGCTACGCTGGGCGGCTTCTGTACTCTGCAAGCCTGGTGGAGCGTGCAGCC 1751
1680 GACCTGCGCTACGCTGGGCGGCTTCTGTGTACTCTGCAAGCCTGGTGGAGCGTGCAGCC 1739
1752 GTCGCGGTGCTCAGCGCGGCCAACCCCGGCGCGGCTCTTACCTGACGAGATGCCCTAT 1811
1740 GTCGCGGTGCTCAGCGCGGCCAACCCCGGCGCGGCTCTTACCTGACGAGATGCCCTAT 1799
1812 CCGTGTATGTGAACGAGCTGTTCTCGTGTGCTGAGCGGCTGCTGCGGCTCTTCTG 1871
1800 CCGTGTATGTGAACGAGCTGTTCTCGTGTGCTGAGCGGCTGCTGCGGCTCTTCTG 1859
1872 AGCTGGCCTTGGATCTACTCCGTGACACTGACAGTGAAGSCGCTGGTGGGAGAGGAG 1931
1860 AGCTGGCCTTGGATCTACTCCGTGACACTGACAGTGAAGSCGCTGGTGGGAGAGGAG 1919
1932 ACSCGCTGCGGACACCAATGCGCGCATTGGGCTCAGCGCGGCTGCTGCGCTAGGC 1991
1920 ACSCGCTGCGGACACCAATGCGCGCATTGGGCTCAGCGCGGCTGCTGCGCTAGGC 1979
1992 TGGTTCTCAGCTGCTCGGCGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTGCTGCTC 2051
1980 TGGTTCTCAGCTGCTCGGCGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTGCTGCTC 2039
2052 AAGCTGGGGACATCTCTCCCTACAGCCACCGGGCGTGTCTTCTCTTGGCAGCC 2111

2040 AAGCTGGGGACATCTCTCCCTACAGCCACCGGGCGTGTCTTCTGCTTGGCAGCC 2099
2112 TTCGCGGTGGCCACAGGTGACCCAGAGCTTCTGTCTCAGCGCCTTCTTCTCCCGCGCAAC 2171
2100 TTCGCGGTGGCCACAGGTGACCCAGAGCTTCTGTCTCAGCGCCTTCTTCTCCCGCGCAAC 2159
2172 CTGGCTTGGCGCTTGGCGGCTTGGGCTTCTTCTCCCTCTACTCTGCTTCTGCTGCTGTGT 2231
2160 CTGGCTTGGCGCTTGGCGGCTTGGGCTTCTTCTCCCTCTACTCTGCTTCTGCTGCTGTGT 2219
2232 GTGGCTTGGCGGACCCGCTGCTCCGCGGTGGCGGCTGGCGCGAGCCTGCTGTGCTGCC 2291
2220 GTGGCTTGGCGGACCCGCTGCTCCGCGGTGGCGGCTGGCGCGAGCCTGCTGTGCTGCC 2279
2292 GTGGCCTTGGCTTGGCTTGGCGAGCCTGGCTTGTCTGGAGAGCAAGGCGAGGCGCG 2351
2280 GTGGCCTTGGCTTGGCTTGGCGAGCCTGGCTTGTCTGGAGAGCAGGCGAGGCGCG 2339
2352 CAGTGGCAAAAGTGGGACCCGGCTACGGCAGAGCTCTTCAAGCTGGCGCAGGTCTCT 2411
2340 CAGTGGCAAAAGTGGGACCCGGCTACGGCAGAGCTCTTCAAGCTGGCGCAGGTCTCT 2399
2412 GGCCTTCTGCTGTGACCGCGCGCTTCTACGGCCTCGCCACCTGGTACCTTGGAGCTGTG 2471
2400 GGCCTTCTGCTGTGACCGCGCGCTTCTACGGCCTCGCCACCTGGTACCTTGGAGCTGTG 2459
2472 TGCCCAAGGCGATGACGGGATCCCTGAAACATGGAAATTTTCTTTCGGAGAGCTACTGG 2531
2460 TGCCCAAGGCGATGACGGGATCCCTGAAACATGGAAATTTTCTTTCGGAGAGCTACTGG 2519
2532 TGGGACCTTGGCGCCCGCCCAAGAGTCCAGCCCTTGGCCACCCCGCTGGACCCAAAGTGG 2591
2520 TGGGACCTTGGCGCCCGCCCAAGAGTCCAGCCCTTGGCCACCCCGCTGGACCCAAAGTGG 2579
2592 CTGGTAGAAGAGGACACCGCCCGCGCTGAGTCTTGGGCTATCCGCTTCCGACCTTGGAGAAG 2651
2580 CTGGTAGAAGAGGACACCGCCCGCGCTGAGTCTTGGGCTATCCGCTTCCGAGCTTGGAGAAG 2639
2652 CGCTTTCCTGGAGGCCCGCAGCAGCCCTTGGCGGCTCAGCCTGAGCTTCTACCAAGGC 2711
2640 CGCTTTCCTGGAGGCCCGCAGCAGCCCTTGGCGGCTCAGCCTGAGCTTCTACCAAGGC 2699
2712 CACATCACCGCTTCTTGGGCGCACACCGGGCGCGCAAGACACCAACCTGCTCACTCTG 2771
2700 CACATCACCGCTTCTTGGGCGCACACCGGGCGCGCAAGACACCAACCTGCTCACTCTG 2759
2772 AGTGGCCTTCTTCCACCCAGTGTGGCTCTGCTTCTCATCTTGGGCCACAGCCTCCGCTCC 2831
2760 AGTGGCCTTCTTCCACCCAGTGTGGCTCTGCTTCTCATCTTGGGCCACAGCCTCCGCTCC 2819
2832 AGCATGGCGCCTATCCGCGCCACCTGGGCTGCTGTCTCTCATCAACGTGCTGTTGAC 2891
2820 AGCATGGCGCCTATCCGCGCCACCTGGGCTGCTGTCTCTCAGTACAACTGCTGTTGAC 2879
2892 ATGCTGACCGTGGACGAGCAGCTTGGTCTTATGGCGGCTGAGGCTGAGTGGCGCT 2951
2880 ATGCTGACCGTGGACGAGCAGCTTGGTCTTATGGCGGCTGAGGCTGAGTGGCGCT 2939
2952 GTAGTGGGCGCCGAGCAGGACCGCTGCTGCTGACGAGTGTGGGGCTGGTCTCCAAAGCAGAGT 3011
2940 GTAGTGGGCGCCGAGCAGGACCGCTGCTGCTGACGAGTGTGGGGCTGGTCTCCAAAGCAGAGT 2999
3012 GTGACAGCTCGCCACCTCTCTGTGGGATGCAACGGAAGCTGTCTGCTGGCCATTTGCCCTTT 3071
3000 GTGACAGCTCGCCACCTCTCTGTGGGATGCAACGGAAGCTGTCTGCTGGCCATTTGCCCTTT 3059
3072 GTGGGCGGCTCCCAAGTGTGTTATCCCTGGACGAGCTACCGGCTGGCGTGGATCTGCTTCC 3131
3060 GTGGGCGGCTCCCAAGTGTGTTATCCCTGGACGAGCTACCGGCTGGCGTGGATCTGCTTCC 3119
3132 CGCCGCGGTATTTGGGAGCTGCTGCTCAAAATACCGAAGAGGTGCGACGCTGATCTCTCC 3191

Db 3120 CGCCGGGTATTGGAGTGTGCTCAATAACGAGAAAGTGCAGCGTGAATCTCTCC 3179
QY 3192 ACCCAACCACTGGATGAGCAGAGCTGTGGAGACCGTGTGGCTGTGTGGCAGGTGGC 3251
Db 3180 ACCCAACCACTGGATGAGCAGAGCTGTGGAGACCGTGTGGCCGTGGTGGCAGGTGGC 3239
QY 3252 CGCTTGTGCTGTGGCTTCCCACTCTTCTCTGGCCGCTCACTGGGCTCGGGCTACTAC 3311
Db 3240 CGCTTGTGCTGTGGATCCCACTCTTCTCTGGCCGCTCACTGGGCTCGGGCTACTAC 3299
QY 3312 CTGACGCTGGTGAAGCGCCGCTGCCCTGACCAACCAATGAGAAGGCTGACACTGACATG 3371
Db 3300 CTGACGCTGGTGAAGCGCCGCTGCCCTGACCAACCAATGAGAAGGCTGACACTGACATG 3359
QY 3372 GAGGCACTGTGGACACCAAGGAGGAAAGAAATGGCAGCCAGGGCAGAGAGTGGC 3431
Db 3360 GAGGCACTGTGGACACCAAGGAGGAAAGAAATGGCAGCCAGGGCAGAGAGTGGC 3419
QY 3432 ACTCTCAGCTGTGGCCCTGTGTACAGCACTGGGTGCCGGGGCA CGGCTGTGGAGAG 3491
Db 3420 ACTCTCAGCTGTGGCCCTGTGTACAGCACTGGGTGCCGGGGCACGGCTGTGGAGAG 3479
QY 3492 CTGCCACACGAGCTGGTGTGGTGTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCC 3551
Db 3480 CTGCCACACGAGCTGGTGTGGTGTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCC 3539
QY 3552 ACACCTCTTCCGAGACTAGACACGGGCTGGCGAGCTGAGGCTCACTGGGCTACGGGATC 3611
Db 3540 ACACCTCTTCCGAGACTAGACACGGGCTGGCGAGCTGAGGCTCACTGGGCTACGGGATC 3599
QY 3612 TCCGACACAGCCTCGAGAGATCTTCTGAAAGTGTGGAGAGTGTGCTGCGGACACA 3671
Db 3600 TCCGACACAGCCTCGAGAGATCTTCTGAAAGTGTGGAGAGTGTGCTGCGGACACA 3659
QY 3672 GATATGAGAGTGGCAGCTGGGCGGAGCACCTATGCAAGGCAATGCTGGGCTAGACGTA 3731
Db 3660 GATATGAGAGTGGCAGCTGGGCGGAGCACCTATGCAAGGCAATGCTGGGCTAGACGTA 3719
QY 3732 ACCCTGCGGCTCAAGATCGGCCACAGAGACAGCGCTGGAGAACGGGAAACCAAGCTGGG 3791
Db 3720 ACCCTGCGGCTCAAGATCGGCCACAGAGACAGCGCTGGAGAACGGGAAACCAAGCTGGG 3779
QY 3792 TCAGCCCCAGAGACTGA CAGGGCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGG 3851
Db 3780 TCAGCCCCAGAGACTGA CAGGGCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGG 3839
QY 3852 GCACCTGACCGCCAGAGCTCCAGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCCG 3911
Db 3840 GCACCTGACCGCCAGAGCTCCAGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCCG 3899
QY 3912 AGCCGCGCGGCTTTCGCCCAGATCGTGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3971
Db 3900 AGCCGCGCGGCTTTCGCCCAGATCGTGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3959
QY 3972 GTGTTACGCTCATCGTGCCTCTTTTCGGGCACTAACCGGCTCTCGGCTCAGTCCCAAC 4031
Db 3960 GTGTTACGCTCATCGTGCCTCTTTTCGGGCACTAACCGGCTCTCGGCTCAGTCCCAAC 4019
QY 4032 ATGTACGCTGTGCTAGGTGCTCTTCTTCAAGTGAAGAGCGCCAGGGGACCTTGAGCTGCC 4091
Db 4020 ATGTACGCTGTGCTAGGTGCTCTTCTTCAAGTGAAGAGCGCCAGGGGACCTTGAGCTGCC 4079
QY 4092 CGGCTGTCTGAGGCGCTCTGTCAGGAGGACGAGCTGAGGAGCGCCCAAGTGCAGCATAGC 4151
Db 4080 CGGCTGTCTGAGGCGCTCTGTCAGGAGGACGAGCTGAGGAGCGCCCAAGTGCAGCATAGC 4139
QY 4152 TCCCA CAGGTTCTCGGCA CCAAGAGTCTCTGCTGAAGTGGCAAGGTCTTGGCCAGTGGC 4211
Db 4140 TCCCA CAGGTTCTCGGCA CCAAGAGTCTCTGCTGAAGTGGCCNAGGTCTTGGCCAGTGGC 4199
QY 4212 AACTGGACCCAGAGTCTCCATTCGCCAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCG 4271
Db 4200 AACTGGACCCAGAGTCTCCATTCGCCAGCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCG 4259

QY 4272 CTGCTGCCCGACTGCCCGGCTGCACTGGTGGTCCCTCCGCCCCAGGCACTGACCGGC 4331
Db 4260 CTGCTGCCCGACTGCCCGGCTGCACTGGTGGTCCCTCCGCCCCAGGCACTGACCGGC 4319
QY 4332 TCTGGGGAAGTGGTTCAAGAACTGTGACAGCCGGAAACCTGTGTGACTTCTCTGTCAAGACC 4391
Db 4320 TCTGGGGAAGTGGTTCAAGAACTGTGACAGCCGGAAACCTGTGTGACTTCTCTGTCAAGACC 4379
QY 4392 TACCCGCGCTTGGTGGCGCAGCGCTGAAGACTAAGAGTGGTGAATGAGGTCAAGTAC 4451
Db 4380 TACCCGCGCTTGGTGGCGCAGCGCTGAAGACTAAGAGTGGTGAATGAGGTCAAGTAC 4439
QY 4452 GGAGGCTTCTCGCTGGGGGGCGGAGACCCAGGCTCGCCCTCGGGCCAAAGAGTTGGGCGGC 4511
Db 4440 GGAGGCTTCTCGCTGGGGGGCGGAGACCCAGGCTCGCCCTCGGGCCAAAGAGTTGGGCGGC 4499
QY 4512 TCAGTGGAGAGTGTGTGGGCGCTGCTGAGTCCCTGCTGGCGGGGCTCTGACCGGTGC 4571
Db 4500 TCAGTGGAGAGTGTGTGGGCGCTGCTGAGTCCCTGCTGGCGGGGCTCTGACCGGTGC 4559
QY 4572 CTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGG 4631
Db 4560 CTGAAAAACCTCACAGCCTGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGG 4619
QY 4632 TTCAACAACAAGGCTGGCACTTCCATGGTGGCTTTTGTCAACCCAGGCGAGCAACCGCAATC 4691
Db 4620 TTCAACAACAAGGCTGGCACTTCCATGGTGGCTTTTGTCAACCCAGGCGAGCAACCGCAATC 4679
QY 4692 CTCGGTGTCACTTCGCCCCCAGGGCCCGCCACAGCCACAGCATCACACACTCAAC 4751
Db 4680 CTCGGTGTCACTTCGCCCCCAGGGCCCGCCACAGCCACAGCATCACACACTCAAC 4739
QY 4752 CACCCCTTGAACCTCACCAAGGAGCGTGTGAGGCTGCACGTAGTGGCTCTCCGCTGG 4811
Db 4740 CACCCCTTGAACCTCACCAAGGAGCGTGTGAGGCTGCACGTAGTGGCTCTCCGCTGG 4799
QY 4812 GACGCTCTGCTCTCATCTGTGTGCTTTTGCCCATGCTTTGTGCCCGGCGAGCTTCACT 4871
Db 4800 GACGCTCTGCTCTCATCTGTGTGCTTTTGCCCATGCTTTGTGCCCGGCGAGCTTCACT 4859
QY 4872 CTGTGCTCATATGAGGAGCGAGTCA CCGAGCCCAAGCACTGCACTCATGGGGGCGCTG 4931
Db 4860 CTGTGCTCATATGAGGAGCGAGTCA CCGAGCCCAAGCACTGCACTCATGGGGGCGCTG 4919
QY 4932 TCCCCCA CCGCTACTGCTGTGGCAACCTTTCTCTGGGACATGTGTAACTACTGTGTGCA 4991
Db 4920 TCCCCCA CCGCTACTGCTGTGGCAACCTTTCTCTGGGACATGTGTAACTACTGTGTGCA 4979
QY 4992 GCATGCATCGTGTGCTCATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGCC 5051
Db 4980 GCATGCATCGTGTGCTCATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGCC 5039
QY 5052 AACCTGCTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATACACACCGCTCATG 5111
Db 5040 AACCTGCTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATACACACCGCTCATG 5099
QY 5112 TACCCAGGCTCTTCTTCTTCTCGTGGCCAGCA CAGGCTATGTGGTGTCACTGCTGCATA 5171
Db 5100 TACCCAGGCTCTTCTTCTTCTCGTGGCCAGCA CAGGCTATGTGGTGTCACTGCTGCATA 5159
QY 5172 AACCTCTTTTATGGCATCAATGGAAGCATGGCCACCTTTTGTGCTTGTAGCTCTTCTCTCAT 5231
Db 5160 AACCTCTTTTATGGCATCAATGGAAGCATGGCCACCTTTTGTGCTTGTAGCTCTTCTCTCAT 5219
QY 5232 CAGAAGCTGCAGGAGGTGAGCCGGATCTTGAAA CAGGCTTTCCTTATCTTCCCCACCTTC 5291
Db 5220 CAGAAGCTGCAGGAGGTGAGCCGGATCTTGAAA CAGGCTTTCCTTATCTTCCCCACCTTC 5279
QY 5292 TGCTTGGCGCGGGGCTCATTTGACATGGTGGGAAAC CAGGCGCATGGCTGTGCTTGGAG 5351
Db 5280 TGCTTGGCGCGGGGCTCATTTGACATGGTGGGAAAC CAGGCGCATGGCTGTGCTTGGAG 5339

Qy	5352	CGCTTGGGAGACAGCGAGTTCCAGTCACCCCTGCGCTGGGAGTGTGCGCAAGAACCTC	5411
Db	5340	CGCTTGGGAGACAGCGAGTTCCAGTCACCCCTGCGCTGGGAGTGTGCGCAAGAACCTC	5399
Qy	5412	TTGGCCATGTTGATACAGGGGCCCTCTTCTTCTCTTCACTACTCTGCTGCAGCACCGA	5471
Db	5400	TTGGCCATGTTGATACAGGGGCCCTCTTCTTCTCTTCACTACTCTGCTGCAGCACCGA	5459
Qy	5472	AGGCCAACTCTTGCCACACAGCCAGGGTGAGTCTCTGCCACTCTCTGGGAGAGGAGACGAG	5531
Db	5460	AGCCAACTCTTGCCACACAGCCAGGGTGAGTCTCTGCCACTCTCTGGGAGAGGAGACGAG	5519
Qy	5532	GATGTAGCCCGTGAACCGGAGCGGGTGGTCCAAAGGAGCCACCCAGGGGGATGTTTGGTG	5591
Db	5520	GATGTAGCCCGTGAACCGGAGCGGGTGGTCCAAAGGAGCCACCCAGGGGGATGTTTGGTG	5579
Qy	5592	CTGAGGAACTTGACCAAGGTATACCGTGGGACAGAGATGCCAGCTGTTGACCGCTTGTC	5651
Db	5580	CTGAGGAACTTGACCAAGGTATACCGTGGGACAGAGATGCCAGCTGTTGACCGCTTGTC	5639
Qy	5652	CTGGGGATTCGCCCTGCTGAGTGTTTTGGGTGCTGCTGGGTGTGAATGAGCAGCGGAGACG	5711
Db	5640	CTGGGGATTCGCCCTGCTGAGTGTTTTGGGTGCTGCTGGGTGTGAATGAGCAGCGGAGACG	5699
Qy	5712	TCCACGTTTCGATGTGTGAACCGGGGACACATTTGGCCACAGGGGCGAGGCTGTGCTGGCA	5771
Db	5700	TCCACGTTTCGATGTGTGAACCGGGGACACATTTGGCCACAGGGGCGAGGCTGTGCTGGCA	5759
Qy	5772	GGCCACAGCTGGGCCCGGGAACCCAGTGTCTGCGCACCTCAGCATGGGATACTGCCCTCAA	5831
Db	5760	GGCCACAGCTGGGCCCGGGAACCCAGTGTCTGCGCACCTCAGCATGGGATACTGCCCTCAA	5819
Qy	5832	TCCGATGCACTTTTGAAGTGTCTGACGGGCGCGAGCACCTGGAGCTGCTTGCGCGCTG	5891
Db	5820	TCCGATGCACTTTTGAAGTGTCTGACGGGCGCGAGCACCTGGAGCTGCTTGCGCGCTG	5879
Qy	5892	CGCGGTGTCCGGAGGCCACAGTTGCCAGACCGCTGGCTCAGGCCCTGGCGCGCTGCGGA	5951
Db	5880	CGCGGTGTCCGGAGGCCACAGTTGGTTGCCAGACCGCTGGCTCAGGCCCTGGCGCGCTGCGGA	5939
Qy	5952	CTCTCATGTGTACGACAGACCGGCTGCGAGGACCTTACAGCGGAGGGGAACAAACGCAAGCTG	6011
Db	5940	CTCTCATGTGTACGACAGACCGGCTGCGAGGACCTTACAGCGGAGGGGAACAAACGCAAGCTG	5999
Qy	6012	CGACGGCCCTGGCGCTGTTTGGGGACCCAGCCGCTGCTGTTCTTGGACGAGCCGACCA	6071
Db	6000	CGACGGCCCTGGCGCTGTTTGGGGACCCAGCCGCTGCTGTTCTTGGACGAGCCGACCA	6059
Qy	6072	GGCATGACCCACGCGCGCGCTCTCTTTGGAAACAGCCCTTTTGGCCGTGCTGCGGAG	6131
Db	6060	GGCATGACCCACGCGCGCGCTCTCTTTGGAAACAGCCCTTTTGGCCGTGCTGCGGAG	6119
Qy	6132	GGCCGTTTCAGTGATGCTCACTCCCATAGCATGGAGAGTGTGAGCGCTCTGCTCGCGC	6191
Db	6120	GGCCGTTTCAGTGATGCTCACTCCCATAGCATGGAGAGTGTGAGCGCTCTGCTCGCGC	6179
Qy	6192	CTAGCCATCATGTGTAATGGSCGGTTCGCTGCTGGCGAGCCCGCAACATCTCAAGGGC	6251
Db	6180	CTAGCCATCATGTGTAATGGSCGGTTCGCTGCTGGCGAGCCCGCAACATCTCAAGGGC	6239
Qy	6252	AGATTGCGCGGGTCAACACTGACCTTCGGGTGCCCGCGCAAGGTCCCCAGCGGCA	6311
Db	6240	AGATTGCGCGGGTCAACACTGACCTTCGGGTGCCCGCGCAAGGTCCCCAGCGGCA	6299
Qy	6312	CGCGCTTCTGTGGCGGCGAGTTCCCTGGTTCGAGTCTGCGGAGGCACATGAGGCGCG	6371
Db	6300	CGCGCTTCTGTGGCGGCGAGTTCCCTGGGCGGAGCTGCGGAGGCACATGAGGCGCG	6359
Qy	6372	CTGCGCTTTCAGCTGCGCGGAGGCGCTGCGCCCTGCGCGCGCTCTTTGGAGAGCTG	6431
Db	6360	CTGCGCTTTCAGCTGCGCGGAGGCGCTGCGCCCTGCGCGCGCTCTTTGGAGAGCTG	6419
Qy	6432	CGGCTGACCGCGCAGACGACGCGCTGGAGGACTTTTTTCGTGAGCCACAGCATGCTGGAG	6491

[illegible]

1795 TGAGCGGTGCTGCGCTCTTCTCTGACGCTGCGCTGATCTACTCGTGCACACTGACAG 1854
1906 TGAAGGCGGTGCTGCGGGAGAGAGACGCGGCTGCGGGACACCAATGCGGCCCATGAGGGC 1965
1855 TGAAGGCGGTGCTGCGGGAGAGAGACGCGGCTGCGGGACACCAATGCGGCCCATGAGGGC 1914
1966 TCAGCGCGGCTGCTCTGCTGCTAGGCTGCTTCTCTAGCTGCTGCGGCCCTTCTGCTCA 2025
1915 TCAGCGCGGCTGCTCTGCTGCTAGGCTGCTTCTCTAGCTGCTGCGGCCCTTCTGCTCA 1974
2026 GCGCGCGCTGCTGCTGCTGCTCAAGCTGCGGGACATCTCCCTACAGCCACCGCG 2085
1975 GCGCGCGCTGCTGCTGCTCAAGCTGCGGGACATCTCCCTACAGCCACCGCG 2034
2086 GCGTGTCTTCTGCTTCTGCGACGCTTCTGCGGTGCGCAAGGTGACCGAGCTTCTGTC 2145
2035 GCGTGTCTTCTGCTTCTGCGACGCTTCTGCGGTGCGCAAGGTGACCGAGCTTCTGTC 2094
2146 TCAGCGCTTCTTCTCCCGCGCAACCTGCTGCGGCTGCGGCGCTGCGGCGCTTCTTCT 2205
2095 TCAGCGCTTCTTCTCCCGCGCAACCTGCTGCGGCTTCTGCGGCGCTTCTTCTTCT 2154
2206 CCCTCTACTGCTGCTGCTGCTGCTGCTGCGGACCGGCTGCGCGCGGCTGCGG 2265
2155 CCCTCTACTGCTGCTGCTGCTGCTGCGGACCGGCTGCGCGCGGCTGCGG 2214
2266 GCGTGTGCGGAGCTGCTGCTGCGCGCTGCTGCTGCGGCTGCGGAGCTGCTGCTC 2325
2215 GCGTGTGCGGAGCTGCTGCTGCGCGCTGCTGCTGCGGCTGCGGAGCTGCTGCTC 2274
2326 TGCTGAGGAGCAGGCGGCGGCGGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2385
2275 TGCTGAGGAGCAGGCGGCGGCGGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2334
2386 ACCTCTTCTGCTGCGGCGGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2445
2335 ACCTCTTCTGCTGCGGCGGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2394
2446 TCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2505
2395 TCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2454
2506 ATTTCTCTTCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2565
2455 ATTTCTCTTCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2514
2566 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2625
2515 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2574
2626 GCGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2685
2575 GCGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2634
2686 GCGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2745
2635 GCGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2694
2746 GCGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2805
2695 GCGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2754
2806 TCATCTGCGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2865
2755 TCATCTGCGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2814
2866 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2925
2815 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2874
2926 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2985

2875 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2934
2986 ATGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3045
2935 ATGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 2994
3046 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3105
2995 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3054
3106 CTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3165
3055 CTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3114
3166 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3225
3115 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3174
3226 ACCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3285
3175 ACCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3234
3286 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3345
3235 GCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3294
3346 CCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3405
3295 CCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3354
3406 ATGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3465
3355 ATGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3414
3466 TGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3525
3415 TGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3474
3526 CCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3585
3475 CCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3534
3586 AGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3645
3535 AGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3594
3646 TGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3705
3595 TGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3654
3706 GCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3765
3655 GCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3714
3766 CCGTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3825
3715 CCGTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3774
3826 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3885
3775 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3834
3886 TCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3945
3835 TCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3894
3946 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 4005
3895 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 3954
4006 ACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 4065
3955 ACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGAGCTGCGGCGCTTCTGCGG 4014

QY	4066	ACGCCCCAGGGGACCCCTGGAACGTGCCCGGCTCTCGAGCGCTGTGAGGAGGAGGAC	4125
Db	4015	ACGCCCCAGGGGACCCCTGGAACGTGCCCGGCTCTCGAGCGCTGTGAGGAGGAGGAC	4074
QY	4126	TGGAGAGCCCCAGTGCAGCATAGCTCCACAGSTTCTCGGCAACAGAGTTCTGTG	4185
Db	4075	TGGAGAGCCCCAGTGCAGCATAGCTCCACAGSTTCTCGGCAACAGAGTTCTGTG	4134
QY	4186	AAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCAGGCTGCC	4245
Db	4135	AAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCAGGCTGCC	4194
QY	4246	AGTGTAGCAGCCCCGGTCCCGGCGCCTGCTGCCCCGATGTCGCCGCTGCAGCTGGTGTG	4305
Db	4195	AGTGTAGCAGCCCCGGTCCCGGCGCCTGCTGCCCCGATGTCGCCGCTGCAGCTGGTGTG	4254
QY	4306	CCCCCTCCGCCAGGACGTGACCCGGCTCTGGGGAAGTGGTTGAGAACCTGACAGGCCGA	4365
Db	4255	CCCCCTCCGCCAGGACGTGACCCGGCTCTGGGGAAGTGGTTGAGAACCTGACAGGCCGA	4314
QY	4366	ACCTGTCTGACTTCTGTGTCGAAGACTTACCCGCGCTGCTGGGGAAGTGGTTGAGAACCTG	4425
Db	4315	ACCTGTCTGACTTCTGTGTCGAAGACTTACCCGCGCTGCTGGGGAAGTGGTTGAGAACCTG	4374
QY	4426	AGAAAGTGGGTGAATCAGAGTTCAGGTACGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCC	4485
Db	4375	AGAAAGTGGGTGAATCAGAGTTCAGGTACGAGGCTTCTCGCTGGGGGGCCGAGACCCAGGCC	4434
QY	4486	TGCCCTCCGGCCAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGCGCTGCTGAGTCCCC	4545
Db	4435	TGCCCTCCGGCCAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGCGCTGCTGAGTCCCC	4494
QY	4546	TGCCCTGGGGGGCCCTCGACCGTGTCTGAAACCTCACAGCCTGGGCTCACAGCCTGG	4605
Db	4495	TGCCCTGGGGGGCCCTCGACCGTGTCTGAAACCTCACAGCCTGGGCTCACAGCCTGG	4554
QY	4606	ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGGGCACTCCATGGTGGCCT	4665
Db	4555	ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAAGGCTGGGCACTCCATGGTGGCCT	4614
QY	4666	TGTCTAACCGAGCCAGCAAGCAATCTCCGTGTCTCAGCTGCCCGCCAGGCGCGGCCGCC	4725
Db	4615	TGTCTAACCGAGCCAGCAAGCAATCTCCGTGTCTCAGCTGCCCGCCAGGCGCGGCCGCC	4674
QY	4726	ACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACAGCCTGGGCTCACAGCCTGG	4785
Db	4675	ACGCCACAGCATCACACACTCAACACCCCTTGAACCTCACAGCCTGGGCTCACAGCCTGG	4734
QY	4786	AGGCTGCACTGATGGGCTCTCGGTGGACGTCCTCGTCTCATCTGTGTGCTTTTGCCA	4845
Db	4735	AGGCTGCACTGATGGGCTCTCGGTGGACGTCCTCGTCTCATCTGTGTGCTTTTGCCA	4794
QY	4846	TGTCTTTGTCCCGGCCAGTCTTCACTTTGTCTCTCAATTTGAGAGGAGTCAACCGAGCCA	4905
Db	4795	TGTCTTTGTCCCGGCCAGTCTTCACTTTGTCTCTCAATTTGAGAGGAGTCAACCGAGCCA	4854
QY	4906	AGCACTGCACTCATGGGGGGCTGTCCCGCAACCTCTACTGTGGCTGGCAACTTCTCT	4965
Db	4855	AGCACTGCACTCATGGGGGGCTGTCCCGCAACCTCTACTGTGGCTGGCAACTTCTCT	4914
QY	4966	GGGACATGTGTAACTTGTGTGGCAGCATGCATGCTGGTGTCTCATCTTTCTGGCCTTCC	5025
Db	4915	GGGACATGTGTAACTTGTGTGGCAGCATGCATGCTGGTGTCTCATCTTTCTGGCCTTCC	4974
QY	5026	AGCAGAGGGCATATGTGGCCCCCTGCACAACTGCTGTCTCTCTGTGTGTGTAATACTGT	5085
Db	4975	AGCAGAGGGCATATATGTGGCCCCCTGCACAACTGCTGTCTCTCTGTGTGTGTAATACTGT	5034
QY	5086	ATGCTGTGTGCATCACACCGCTCATGTATACCGGCTCTCTCTCTCTCTCTCTCTCTCTCT	5145
Db	5035	ATGCTGTGTGCATCACACCGCTCATGTATACCGGCTCTCTCTCTCTCTCTCTCTCTCTCT	5094

QY	5146	CAGCTATGTGTGTCTCACTGTCAATAAACCTCTTTTATTGGCATCAATGGAAGCATGGCCA	5205
Db	5095	CAGCTATGTGTGTCTCACTGTCAATAAACCTCTTTTATTGGCATCAATGGAAGCATGGCCA	5154
QY	5206	CCTTTGTGCTTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGAGTGTAGCCCGGATCTTGAAC	5265
Db	5155	CCTTTGTGCTTTGAGCTCTTCTCTGATCAGAAGCTGCAGAGGAGTGTAGCCCGGATCTTGAAC	5214
QY	5266	AGGTCTTCTCTTATCTTTCCCCCACTTCTGTGGGGCCGGGGCTATTGACATGGTGGGA	5325
Db	5215	AGGTCTTCTCTTATCTTTCCCCCACTTCTGTGGGGCCGGGGCTATTGACATGGTGGGA	5274
QY	5326	ACCAAGCCATGGCTGATGCTTTGAGCGCTTTGGAGACAGGACAGTTCAGTCAACCCCTGTC	5385
Db	5275	ACCAAGCCATGGCTGATGCTTTGAGCGCTTTGGAGACAGGACAGTTCAGTCAACCCCTGTC	5334
QY	5386	GCTGGGAGGTGGTGGGCAAGAACCTCTTTGGGCATGGTGTATACAGGGGGCCCTCTTCTTCT	5445
Db	5335	GCTGGGAGGTGGTGGGCAAGAACCTCTTTGGGCATGGTGTATACAGGGGGCCCTCTTCTTCT	5394
QY	5446	TCCTTCACTACTTGTCTGAGCACCGAAGCCAACTCTGCCACAGCCAGGCTGAGGTCTTC	5505
Db	5395	TCCTTCACTACTTGTCTGAGCACCGAAGCCAACTCTGCCACAGCCAGGCTGAGGTCTTC	5454
QY	5506	TGCCACTCTCTGGGAGAGGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGTGCTCAAG	5565
Db	5455	TGCCACTCTCTGGGAGAGGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGTGCTCAAG	5514
QY	5566	GAGCCACCCAGGGGATGTGTGTGTGTGAGGAACTTGAACCAAGGTATACCGTGGGGCAGA	5625
Db	5515	GAGCCACCCAGGGGATGTGTGTGTGTGAGGAACTTGAACCAAGGTATACCGTGGGGCAGA	5574
QY	5626	GGATGCCAGCTGTGTGACCGCTTGTCCCTGGGGATTTCCCTTGGTGTGTTTGGGCTGTC	5685
Db	5575	GGATGCCAGCTGTGTGACCGCTTGTCCCTGGGGATTTCCCTTGGTGTGTTTGGGCTGTC	5634
QY	5686	TGGGTGTGAATGGAGCAGGGAAGACGTCACCGTTTTCGATGTGTGACGGGGGACACATTTGG	5745
Db	5635	TGGGTGTGAATGGAGCAGGGAAGACGTCACCGTTTTCGATGTGTGACGGGGGACACATTTGG	5694
QY	5746	CCACAGGGGGCAGGCTGTGTGTGAGGCAACAGGCTGGCCGGGAAACCGAGTGTGTGGC	5805
Db	5695	CCACAGGGGGCAGGCTGTGTGTGAGGCAACAGGCTGGCCGGGAAACCGAGTGTGTGGC	5754
QY	5806	ACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGCG	5865
Db	5755	ACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGCG	5814
QY	5866	AGCACTTGGAGCTGCTTTGCGCGCTTGGCGGCTGTCCGGAGGGCCAGGTTGCCAGACCG	5925
Db	5815	AGCACTTGGAGCTGCTTTGCGCGCTTGGCGGCTGTCCGGAGGGCCAGGTTGCCAGACCG	5874
QY	5926	CTGGCTCAGGCTGGCGGCTTGGGACTCTCATGTGTACGAGACCGGCTGTGAGGCACT	5985
Db	5875	CTGGCTCAGGCTGGCGGCTTGGGACTCTCATGTGTACGAGACCGGCTGTGAGGCACT	5934
QY	5986	ACAGCGAGGGAACAAACGCAAGCTGCGGACCGGCTGTGGGCTGTGGGACCCAGCGG	6045
Db	5935	ACAGCGAGGGAACAAACGCAAGCTGCGGACCGGCTGTGGGCTGTGGGACCCAGCGG	5994
QY	6046	TGGTGTCTTCTGGAACGAGCCGACACAGGCAATGGACCCCGCGCGGGCGCTTCTCTTGA	6105
Db	5995	TGGTGTCTTCTGGAACGAGCCGACACAGGCAATGGACCCCGCGCGGGCGCTTCTCTTGA	6054
QY	6106	ACAGCTTTTGGCGGTGGTGGGAGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6165
Db	6055	ACAGCTTTTGGCGGTGGTGGGAGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6114
QY	6166	AGGAGTGTGAAGCGCTCTGTCTCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6225
Db	6115	AGGAGTGTGAAGCGCTCTGTCTCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6174
QY	6226	TGGGACGCGCGCAACATCTCAAGGGCAGATTTCGGGGGGGTTCACACACTGACCTTGGCGG	6285

```
Db      6175  TGGGAGCCCGCAACATCTCAAGGGCAGATTGCGGGCGGGTCACACATGACCTCGGGG 6234
Qy      6286  TGCCCGCGCAAGTCCAGCGCGACGGCGCTTGTGGCGCGCGAGTTCCTCGGGTCGG 6345
Db      6235  TGCCCGCGCAAGTCCAGCGCGACGGCGCTTGTGGCGCGCGAGTTCCTCGGGTCGG 6294
Qy      6346  AGCTGCGCGAGGACATGGAAGGCGCTGCGCTTCCAGCTGCGCGCGGAGGGCGCTGCG 6405
Db      6295  AGCTGCGCGAGGACATGGAAGGCGCTGCGCTTCCAGCTGCGCGCGGAGGGCGCTGCG 6354
Qy      6406  CCCTGGCGCGCGCTTGTGGAGAGCTGCGCGGTGCAACGCGCGAGACGCGCGTGGAGACT 6465
Db      6355  CCCTGGCGCGCGCTTGTGGAGAGCTGCGCGGTGCAACGCGCGAGACGCGCGTGGAGACT 6414
Qy      6466  TTTCCTGGAGCCAGACATGCTGGAGGAGTATTTCTGTACTTCTTCCAGGACGAGGGGA 6525
Db      6415  TTTCCTGGAGCCAGACATGCTGGAGGAGTATTTCTGTACTTCTTCCAGGACGAGGGGA 6474
Qy      6526  AGGACGAGGACACCGAAGACGAGAGGAGGAGGAGTGGAGTGGACCCCGCGCAGGCC 6585
Db      6475  AGGACGAGGACACCGAAGACGAGAGGAGGAGGAGTGGAGTGGACCCCGCGCAGGCC 6534
Qy      6586  TGACGACACCCCAACCGGCTCAGCCAGTTCTCTCGATGACCCCTAGCACTGCGGAGACTGTGC 6645
Db      6535  TGACGACACCCCAACCGGCTCAGCCAGTTCTCTCGATGACCCCTAGCACTGCGGAGACTGTGC 6594
Qy      6646  TCTGAGCTCTCCCTCCCTCGCGGGCCCGCGGGAGGCCCTGGGAAATGGCAAGGGCAAGGTA 6705
Db      6595  TCTGAGCTCTCCCTCCCTCGCGGGCCCGCGGGAGGCCCTGGGAAATGGCAAGGGCAAGGTA 6654
Qy      6706  GAGTGCCCTAGAGCCCTGGACTCAGCTGCGAGGCGGCTGCTGCGGAGGAGGAGGAGGTA 6765
Db      6655  GAGTGCCCTAGAGCCCTGGACTCAGCTGCGAGGCGGCTGCTGCGGAGGAGGAGGAGGTA 6714
Qy      6766  GAGAGGCTGGAGAGAGCGCTGCTGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGTA 6804
Db      6715  GAGAGGCTGGAGAGAGCGCTGCTGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGTA 6753
```

```
RESULT 4
AX429481 LOCUS AX429481 7795 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 8 from Patent WO0234903.
ACCESSION AX429481
VERSION AX429481.1 GI:21540776
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Arnould-Reguigne, I., Deneffe, P., Chimini, G., Duverger, N.,
Osorio, Y. F., Prades, C. and Rosier, M. F.
TITLE Nucleic acid generating the abca7 gene, molecules modulating its
activity and therapeutic applications
JOURNAL Patent: WO 0234903-A 8 02-MAY-2002;
AVENTIS PHARMA SA (FR); INST NAT SANTE RECH MED (FR)
FEATURES
source 1..7795
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.6%; Score 6710; DB 6; Length 7795;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6722; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 56 TGCGGAGCCCTGGAGAGCTTCCAGGAACCTGCGCTGTGGGATAAAGGAATGAGGTTTC 1115
Db 1054 TGCTCGCTCTAATCAGAGCTTCCAGGAACCTGCGCTGTGGGATAAAGGAATGAGGTTTC 1113
```

```
Qy      116  AGAAAGGGGAGGAGCTTCCCGAGCCGCAACGACGCTCTTCAGCCCGACCGCTGTGCCT 175
Db      1114  AGAAAGGGGAGGAGCTTCCCGAGCCGCAACGACGCTCTTCAGCCCGACCGCTGTGCCT 1173
Qy      176  GACCTCTCTGTCCCGTCCCGCCAGTCTACCATGGGCTTCTGGAACAGCTGATGCT 235
Db      1174  GACCTCTCTGTCCCGTCCCGCCAGTCTACCATGGGCTTCTGGAACAGCTGATGCT 1233
Qy      236  GCTGCTCTGGAGAAATTTATGATCGCGGAGACAGCCGGTCCAGCTCTGCTGCAATTT 295
Db      1234  GCTGCTCTGGAGAAATTTATGATCGCGGAGACAGCCGGTCCAGCTCTGCTGCAATTT 1293
Qy      296  GCTGTGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 355
Db      1294  GCTGTGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353
Qy      356  GGAAGCACCATGATGCACTTCCAAACAAAGCCACTGCCATCGCGGGCACCGTGCCTG 415
Db      1354  GGAAGCACCATGATGCACTTCCAAACAAAGCCACTGCCATCGCGGGCACCGTGCCTG 1413
Qy      416  GCTCCAGGGTCTCATCTGTAAATGTGAACAAACACTGCTTTCGCGAGCTGACACGGGCGA 475
Db      1414  GCTCCAGGGTCTCATCTGTAAATGTGAACAAACACTGCTTTCGCGAGCTGACACGGGCGA 1473
Qy      476  GGAAGCCGCGGCGCTGAGCAATTTCAACGACTCCCTGGTCTCCCGGCTGCTAGCCGATGC 535
Db      1474  GGAAGCCGCGGCGCTGAGCAATTTCAACGACTCCCTGGTCTCCCGGCTGCTAGCCGATGC 1533
Qy      536  CGCACACTGTCTGGGAGGGGCCAGTCCCAAGGACGCTGGCTGGCTAGGGAAGCTGAT 595
Db      1534  CGCACACTGTCTGGGAGGGGCCAGTCCCAAGGACGCTGGCTGGCTAGGGAAGCTGAT 1593
Qy      596  CGCACACTGTAGGGGCTGACAGGACGCGCCAGCTCAACCAACCAAGCAGTCTCCACT 655
Db      1594  CGCACACTGTAGGGGCTGACAGGACGCGCCAGCTCAACCAACCAAGCAGTCTCCACT 1653
Qy      656  GGAACCAACCACTGCTGATGTGCGGAGCTGCTGAGCTCTGCTGCGCAGCAAGTCCCT 715
Db      1654  GGAACCAACCACTGCTGATGTGCGGAGCTGCTGAGCTCTGCTGCGCAGCAAGTCCCT 1713
Qy      716  GGGGTTGGCACTGGGGCCAAAGCCAGGAGCCCTTGGACAGCTTGTGGAGCCCTGAGGA 775
Db      1714  GGGGTTGGCACTGGGGCCAAAGCCAGGAGCCCTTGGACAGCTTGTGGAGCCCTGAGGA 1773
Qy      776  CTTGGCCCGAGGAGCTCTTGGCGCTGCGAGCTGCTGGAGCTTTCGGGCACTGCTGAGAG 835
Db      1774  CTTGGCCCGAGGAGCTCTTGGCGCTGCGAGCTTTCGGGCACTGCTGAGAGAG 1833
Qy      836  ACCCCGAGGAGACGAGCGGCCCTGAGTGTGCTCAGAGGCCCTCTGCAAGTGTCAAGGG 895
Db      1834  ACCCCGAGGAGACGAGCGGCCCTGAGTGTGCTCAGAGGCCCTCTGCAAGTGTCAAGGG 1893
Qy      896  ACCTAGCAGCAGTAGTGGGCCCTTCCCTCAACTGGTACGAGGCTAGTGAACCTGATGGAGCT 955
Db      1894  ACCTAGCAGCAGTAGTGGGCCCTTCCCTCAACTGGTACGAGGCTAGTGAACCTGATGGAGCT 1953
Qy      956  GGTGGGGCAGGAGCCAGAAATCCGCCCTGCCAGACAGGAGCTGAGCCCCGCTGCTCGGA 1015
Db      1954  GGTGGGGCAGGAGCCAGAAATCCGCCCTGCCAGACAGGAGCTGAGCCCCGCTGCTCGGA 2013
Qy      1016  GCTGATTGGAGCCCTGAGACGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
Db      2014  GCTGATTGGAGCCCTGAGACGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
Qy      1076  TCTGATCTCTCGGAAGCTACTCTTTTGCACACAGATACACCTTTTACCCCGGAAGCTCATGGC 1135
Db      2074  TCTGATCTCTCGGAAGCTACTCTTTTGCACACAGATACACCTTTTACCCCGGAAGCTCATGGC 2133
Qy      1136  CAGGTGTAACCGGACCTTCGAGAGCTCACCCTGCTGAGGAGTGTCCGGGAGGCTGTGGGA 1195
Db      2134  CAGGTGTAACCGGACCTTCGAGAGCTCACCCTGCTGAGGAGTGTCCGGGAGGCTGTGGGA 2193
Qy      1196  GATGCTGGGAGCCCGGATCTTACCTTCATGAACGACAGTTCCAATGTGCGCCATGCTGCA 1255
```

Db	2194	 GATGCTGGGACCCCGGATCTTTCACTTTCACTGAACGACAGTTTCCAATGTGGCCTAGCTGCA	2253
Qy	1256	CGCGCTCTCTGCAGATGCAGGATGAAGAAAGGACGACCCAGACCTCGAGAGCGCGGACCA	1315
Db	2254	CGCGCTCTCTGCAGATGCAGGATGNAAGAGAGAGGACGACCCAGACCTGAGAGCGCGGACCA	2313
Qy	1316	CATGGAGGCCCTTGCGATCTCTTTCTGGACCTTGGAGCGGTGGCTACAGCTGGCAGGACGC	1375
Db	2314	CATGGAGGCCCTTGCGATCTCTTTCTGGACCTTGGAGCGGTGGCTACAGCTGGCAGGACGC	2373
Qy	1376	ACACGCTGATGTGGGGACACTGTGTGGGCACGCTGGGCGGAGTGACGAGATGCTGTGCTTT	1435
Db	2374	ACACGCTGATGTGGGGACACTGTGTGGGCACGCTGGGCGGAGTGACGAGATGCTGTGCTTT	2433
Qy	1436	GGACAAGCTTGGAGCGGACACCTTCAGAGGACAGCCCTGCTGCGGGGCCCTCGCAACTGCT	1495
Db	2434	GGAACAGCTTGGAGCGGACACCTTCAGAGGACAGCCCTGCTGCGGGGCCCTCGCAACTGCT	2493
Qy	1496	CGCGGAACATCGATTTCTGGGCGGCGCTCTTTCTTGGGACCTTGAGGACTCTTTCAGACCC	1555
Db	2494	CGCGGAACATCGATTTCTGGGCGGCGCTCTTTCTTGGGACCTTGAGGACTCTTTCAGACCC	2553
Qy	1556	CACAGAGCACCGAAACCCAGACCTGGGGCCCGGGCCAGCTGCGGATCAAAATCCGCAATGGA	1615
Db	2554	CACAGAGCACCGAAACCCAGACCTGGGGCCCGGGCCAGCTGCGGATCAAAATCCGCAATGGA	2613
Qy	1616	CATTGACGTGGTCACGAGGACCAATAAGATCAGGACAGGTTTTGGGACCCCTGGCCGAC	1675
Db	2614	CATTGACGTGGTCACGAGGACCAATAAGATCAGGACAGGTTTTGGGACCCCTGGCCGAC	2673
Qy	1676	CGCGGACCCCTTGACACCACTGGCGTACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCT	1735
Db	2674	CGCGGACCCCTTGACACCACTGGCGTACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCT	2733
Qy	1736	GGTGGAGCGTGCAGCGTCCGCTGCTCAGCGGGCGCAACCCCGGGCGCGGCTCTTACCT	1795
Db	2734	GGTGGAGCGTGCAGCGTCCGCTGCTCAGCGGGCGCAACCCCGGGCGCGGCTCTTACCT	2793
Qy	1796	GCAGCAGATGCCCTATCCGCTGCTATGTGGACGACGTGTTCTCGCTGTGCTGAGCGGTC	1855
Db	2794	GCAGCAGATGCCCTATCCGCTGCTATGTGGACGACGTGTTCTCGCTGTGCTGAGCGGTC	2853
Qy	1856	GCTGCCGCTTTCTTGACGCTGGCCCTGGATCTACTCGTGACACTGACAGTGAAGCGCGT	1915
Db	2854	GCTGCCGCTTTCTTGACGCTGGCCCTGGATCTACTCGTGACACTGACAGTGAAGCGCGT	2913
Qy	1916	GGTGGGGAGAGAGACGCGCTGCGGGAACACATGCGGCGCATGGGGCTCAGCCGCGC	1975
Db	2914	GGTGGGGAGAGAGACGCGCGTGGCGGACACATGCGGCGCATGGGGCTCAGCCGCGC	2973
Qy	1976	GGTGCTCTGCTAGGCTGGTTCTCAGCTGCGCTCGGGCCCTTCTGCTCAGCGCGCGCT	2035
Db	2974	GGTGCTCTGCTAGGCTGGTTCTCAGCTGCGCTCGGGCCCTTCTGCTCAGCGCGCGACT	3033
Qy	2036	GCTGGTTCTGGTGTCAAGCTGGGGACATCTCCCTACAGCACCCGCGGCGTGTCTTT	2095
Db	3034	GCTGGTTCTGGTGTCAAGCTGGGGACATCTCCCTACAGCACCCGCGGCGTGTCTTT	3093
Qy	2096	CTTGTTCTTGGCAGCCTTCGCGTGGCCACCGGTGAACCCAGAGCTTCTGCTCAGCGCCTT	2155
Db	3094	CTTGTTCTTGGCAGCCTTCGCGTGGCCACCGGTGAACCCAGAGCTTCTGCTCAGCGCCTT	3153
Qy	2156	CTTCTCCGCGCAACCTTGGCTGGGCGCTCGGGCGGCTGGCCCTACTTCTCCCTCTACCT	2215
Db	3154	CTTCTCCGCGCAACCTTGGCTGGGCGCTCGGGCGGCTGGCCCTACTTCTCCCTCTACCT	3213
Qy	2216	GCCTACGTGCTGTGTGGCTTTGGGGGACCGGCTGCGCGGGGTGGCGCGTGGCGCGC	2275
Db	3214	GCCTACGTGCTGTGTGGCTTTGGGGGACCGGCTGCGCGGGGTGGCGCGTGGCGCGC	3273
Qy	2276	GAGCCTGCTGTGCGCGCTTTCGCTTCGAGAGCCTGGCTCTGCTGTGAGGA	2335

Db	3274	GAGCCTGCTGTGCGCCGTGGCCCTTTCGGCTTCGGCTGCGAGAGCGCTGGCTCTGCTGGAGGA	3333
Qy	2336	GCAGGCGAGGGCGCGCAGTGGCAACAGCTGGGCAACCGCGCCTACGGCAGACGCTTTCAG	2395
Db	3334	GCAAGGCGAGGGCGCGCAGTGGCAACAGCTGGGCAACCGCGCCTACGGCAGACGCTTTCAG	3393
Qy	2396	CCTGGCCCAAGTCTCTGGCCCTTCTGCTGTCTGGAGCGGCGCTCTTACGGCCTTCGCACCTG	2455
Db	3394	CCTGGCCCAAGTCTCTGGCCCTTCTGCTGTCTGGAGCGGCGCTCTTACGGCCTTCGCACCTG	3453
Qy	2456	GTACTCTGGAGCTGTGTGCCAGGCCAGTACGGATCCCTTGAACCATGGAAATTTTCTTT	2515
Db	3454	GTACTCTGGAGCTGTGTGCCAGGCCAGTACGGATCCCTTGAACCATGGAAATTTTCTTT	3513
Qy	2516	TCGAGAGAGCTACTGTGTGGAGCCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCACCCC	2575
Db	3514	TCGAGAGAGCTACTGTGTGGAGCCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCACCCC	3573
Qy	2576	GCTGGACCCAAAGTGTGTGTAGAGAGGCACCGCCGGCTGTAGTCTTGGCGTATCCGT	2635
Db	3574	GCTGGACCCAAAGTGTGTGTAGAGAGGCACCGCCGGCTGTAGTCTTGGCGTATCCGT	3633
Qy	2636	TCGAGAGCTGGAGAGCGCTTTCCTTGAAGCCGCGAGCGCAGCCCTGGGGGGCTCAGCCT	2695
Db	3634	TCGAGAGCTGGAGAGCGCTTTCCTTGAAGCCGCGAGCGCAGCCCTGGGGGGCTCAGCCT	3693
Qy	2696	GGACTTCTACACGGCCACATCACCGCCTTCTGGGCCACAAAGGGCCGGCAGAGCCAC	2755
Db	3694	GGACTTCTACACGGCCACATCACCGCCTTCTGGGCCACAAAGGGCCGGCAGAGCCAC	3753
Qy	2756	CACCTCTGCATCTTGTAGTGGCCTTTCCTCCACCCAGTGGTGTCTGCTTTCATCTCTGGG	2815
Db	3754	CACCTCTGCATCTTGTAGTGGCCTTTCCTCCACCCAGTGGTGTCTGCTTTCATCTCTGGG	3813
Qy	2816	CCAGCAGTCCGCTCCAGATGGCGCGCATCTGGGCCCAACCTGGGCGTCTGTCTCTCAGTA	2875
Db	3814	CCAGCAGTCCGCTCCAGATGGCGCGCATCTGGGCCCAACCTGGGCGTCTGTCTCTCAGTA	3873
Qy	2876	CAACGTCTGTTTGACATGTGACCGTGGACGACGCTCTGCTTCTATGGGCGGCTGAA	2935
Db	3874	CAACGTCTGTTTGACATGTGACCGTGGACGACGCTCTGCTTCTATGGGCGGCTGAA	3933
Qy	2936	GGGTCTGAGTGCCTGTAGTGGGCCCGCAGCAGGACCGTCTGCTGTCAGGATGTGGGCT	2995
Db	3934	GGGTCTGAGTGCCTGTAGTGGGCCCGCAGCAGGACCGTCTGCTGTCAGGATGTGGGCT	3993
Qy	2996	GGTCTCAAGCAGAGTGTGAGACTGCGCACCTCTCTGTGTGGATGCAACGGAAGCTGTC	3055
Db	3994	GGTCTCAAGCAGAGTGTGAGACTGCGCACCTCTCTGTGTGGATGCAACGGAAGCTGTC	4053
Qy	3056	CGTGGCCATTTGCTTTGTGGCGGCTCCCAAGTTGTATCTCTGAGCAGGCTACGGCTGG	3115
Db	4054	CGTGGCCATTTGCTTTGTGGCGGCTCCCAAGTTGTATCTCTGAGCAGGCTACGGCTGG	4113
Qy	3116	CGTGGATCTGCTTCCCGCCGGGTATTTGGGAGCTGCTCTCAAATACCGAAGGTGCG	3175
Db	4114	CGTGGATCTGCTTCCCGCCGGGTATTTGGGAGCTGCTCTCAAATACCGAAGGTGCG	4173
Qy	3176	CAGCTGATCTCTTCCACCCACCTCTGATAGGAGAGCTGCTGGGAGACGCTGTGGC	3235
Db	4174	CAGCTGATCTCTTCCACCCACCTCTGATAGGAGAGCTGCTGGGAGACGCTGTGGC	4233
Qy	3236	TGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTCCCACTTCTTCTGCGCGCTCACCT	3295
Db	4234	TGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTCCCACTTCTTCTGCGCGCTCACCT	4293
Qy	3296	GGGCTCCGGCTACTACTGACGCTGTGTGAAGGCCCGCTTCCCTTACCACCAATGAGAA	3355
Db	4294	GGGCTCCGGCTACTACTGACGCTGTGTGAAGGCCCGCTTCCCTTACCACCAATGAGAA	4353
Qy	3356	GGCTGACATGACATGAGAGGCGAGTGTGGACACAGGCGAGGAAAAGAAATGGCAGCCA	3415
Db	4354	GGCTGACATGACATGAGAGGCGAGTGTGGACACAGGCGAGGAAAAGAAATGGCAGCCA	4413

Qy	3416	GGG	CAG	CAG	AGT	TCGG	CACT	CTCT	CAG	CTG	CTGG	CCCT	CTGGT	TAC	GAC	CTGGGT	GCCCGGGC	3475				
Db	4414	GGG	CAG	CAG	AGT	TCGG	CACT	CTCT	CAG	CTG	CTGG	CCCT	CTGGT	TAC	GAC	CTGGGT	GCCCGGGC	4473				
Qy	3476	ACGG	CTGG	TGAG	GAG	CTGCC	ACAC	ACAG	AGCT	TGGT	CTGG	CTGG	CTCC	TAC	ACGGGT	TGCCCA	3535					
Db	4474	ACGG	CTGG	TGAG	GAG	CTGCC	ACAC	ACAG	AGCT	TGGT	CTGG	CTGG	CTCC	TAC	ACGGGT	TGCCCA	4533					
Qy	3536	TGAC	GGC	GAG	CTT	TGCC	CACT	CTCT	CCAG	AGCT	TAG	ACCG	GGCT	GCG	GAG	CTGAGGCT	3595					
Db	4534	TGAC	GGC	GAG	CTT	TGCC	CACT	CTCT	CCAG	AGCT	TAG	ACCG	GGCT	GCG	GAG	CTGAGGCT	4593					
Qy	3596	CAC	TGG	CTT	ACGG	ATCT	CCG	ACAC	CCAG	CCCT	TCG	AGG	AGAT	CTT	CTG	AAAGT	TGGTGAAGGA	3655				
Db	4594	CAC	TGG	CTT	ACGG	ATCT	CCG	ACAC	CCAG	CCCT	TCG	AGG	AGAT	CTT	CTG	AAAGT	TGGTGAAGGA	4653				
Qy	3656	GTG	TG	CTG	CGGA	CA	CAG	ATAT	TG	GAG	AGAT	TGG	CAG	CTG	CGG	GAG	CACTAT	TGCACAGGCAT	3715			
Db	4654	GTG	TG	CTG	CGGA	CA	CAG	ATAT	TG	GAG	AGAT	TGG	CAG	CTG	CGG	GAG	CACTAT	TGCACAGGCAT	4713			
Qy	3716	TGCT	TGG	CC	TAG	ACGT	TAC	CCCT	TG	CGG	CTCA	AGAT	TGC	CGC	ACAG	GAG	ACGCGT	TGAGAA	3775			
Db	4714	TGCT	TGG	CC	TAG	ACGT	TAC	CCCT	TG	CGG	CTCA	AGAT	TGC	CGC	ACAG	GAG	ACGCGT	TGAGAA	4773			
Qy	3776	CGG	GAAC	CAG	CT	TGGG	CT	CAG	CC	CC	CAG	AGAC	TG	ACC	AGGG	CT	CTGGG	CCACAC	3835			
Db	4774	CGG	GAAC	CAG	CT	TGGG	CT	CAG	CC	CC	CAG	AGAC	TG	ACC	AGGG	CT	CTGGG	CCACAC	4833			
Qy	3836	CCG	GGT	TAC	AGG	GT	TGG	CA	CTG	ACC	CGCC	ACG	AGCT	CTC	AGG	CCCT	CTG	CTT	CTCAAGCGCTT	3895		
Db	4834	CCG	GGT	TAC	AGG	GT	TGG	CA	CTG	ACC	CGCC	ACG	AGCT	CTC	AGG	CCCT	CTG	CTT	CTCAAGCGCTT	4893		
Qy	3896	TCT	GCT	T	GCCG	CGC	GAG	CGC	CGG	CGCT	TG	TTC	GCC	CCAG	ATG	TG	TGCT	GCT	CGCCTCTCTT	3955		
Db	4894	TCT	GCT	T	GCCG	CGC	GAG	CGC	CGG	CGCT	TG	TTC	GCC	CCAG	ATG	TG	TGCT	GCT	CGCCTCTCTT	4953		
Qy	3956	TG	TGGG	CT	TGG	CCCT	CTG	TG	TTC	CAG	CTCAT	CTG	TG	CCCT	TCT	TTC	TGGG	CAC	TAC	CCCGGCTCT	4015	
Db	4954	TG	TGGG	CT	TGG	CCCT	CTG	TG	TTC	CAG	CTCAT	CTG	TG	CCCT	TCT	TTC	TGGG	CAC	TAC	CCCGGCTCT	5013	
Qy	4016	GCG	CT	CAG	T	CCCA	CCAT	GTAC	CGT	GTCT	CAG	GTG	TC	CTT	CTT	CAG	TGAG	AC	GC	CCCAAGG	4075	
Db	5014	GCG	CT	CAG	T	CCCA	CCAT	GTAC	CGT	GTCT	CAG	GTG	TC	CTT	CTT	CAG	TGAG	AC	GC	CCCAAGG	5073	
Qy	4076	GGAC	CT	TG	AG	CTGCC	CGG	CTG	CTC	GAG	GGCG	CTG	CTC	ACG	AGG	CAG	CACT	TG	AG	GAGGC	4135	
Db	5074	GGAC	CT	TG	AG	CTGCC	CGG	CTG	CTC	GAG	GGCG	CTG	CTC	ACG	AGG	CAG	CACT	TG	AG	GAGGC	5133	
Qy	4136	CC	CAG	T	GCAG	TAT	AG	CTT	CC	CA	AG	TTCT	CGG	CA	CCAG	AA	GTTC	TCT	GC	TCAAGTGGCCAA	4195	
Db	5134	CC	CAG	T	GCAG	TAT	AG	CTT	CC	CA	AG	TTCT	CGG	CA	CCAG	AA	GTTC	TCT	GC	TCAAGTGGCCAA	5193	
Qy	4196	GGT	CTT	TGG	CCAG	TGG	CAACT	TGG	AC	CC	CCAG	AGT	CTT	CC	AT	CC	CCAG	CCCT	TG	CAGTGT	4255	
Db	5194	GGT	CTT	TGG	CCAG	TGG	CAACT	TGG	AC	CC	CCAG	AGT	CTT	CC	AT	CC	CCAG	CCCT	TG	CAGTGT	5253	
Qy	4256	GCC	CGGT	CCCC	GGCC	CTG	CTG	CCCC	GA	CTG	CC	GGG	TG	CA	GCT	TGGT	GGT	CCCC	CT	CCGCC	4315	
Db	5254	GCC	CGGT	CCCC	GGCC	CTG	CTG	CCCC	GA	CTG	CC	GGG	TG	CA	GCT	TGGT	GGT	CCCC	CT	CCGCC	5313	
Qy	4316	CCAG	G	CAG	T	GAC	CGG	CT	TGGG	AA	G	TG	TT	CA	GA	CT	CA	AG	CGCG	AA	CTG	4375
Db	5314	CCAG	G	CAG	T	GAC	CGG	CT	TGGG	AA	G	TG	TT	CA	GA	CT	CA	AG	CGCG	AA	CTG	5373
Qy	4376	CTT	CT	GGT	CAAG	AC	TAC	CC	CG	CCCT	TGGT	G	CG	C	AG	GG	CT	CA	AG	ACT	4435	
Db	5374	CTT	CT	GGT	CAAG	AC	TAC	CC	CG	CCCT	TGGT	G	CG	C	AG	GG	CT	CA	AG	ACT	4435	
Qy	4436	GAAT	GAG	GT	CAG	GT	CTT	CG	CT	TGG	GGG	GG	CC	G	AG	CA	CCC	AG	CG	CT	4495	
Db	5434	GAAT	GAG	GT	CAG	GT	CTT	CG	CT	TGG	GGG	GG	CC	G	AG	CA	CCC	AG	CG	CT	5493	

Qy	4496	CCAAGAGTTGGGCGCGCTCAGTGGAGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCGCGG	4555
Db	5494	CCAAGAGTTGGGCGCGCTCAGTGGAGAGTTGTGGGCGCTGCTGAGTCCCTGCTGCGCGG	5553
Qy	4556	GGCCCTCGACGGTGTCTGTGAAAACCTCACAGCCTGGGCTCACAGCCTGATGCTCAGGA	4615
Db	5554	GGCCCTCGACGGTGTCTGTGAAAACCTCACAGCCTGGGCTCACAGCCTGATGCTCAGGA	5613
Qy	4616	CAGTCTCAAGATCTGGTTTCAACAAGAGCTGGCACTCCATGGTGGCCCTTTGTCAACCG	4675
Db	5614	CAGTCTCAAGATCTGGTTTCAACAAGAGCTGGCACTCCATGGTGGCCCTTTGTCAACCG	5673
Qy	4676	AGCCAGCAACGCAATCTCCGTGCTCACTGCCCCCAAGGCCCGCCGCGCACGCCACAG	4735
Db	5674	AGCCAGCAACGCAATCTCCGTGCTCACTGCCCCCAAGGCCCGCCGCGCACGCCACAG	5733
Qy	4736	CATCACCACTCAACCAACCCCTTGAACCTTCAACAGGAGCAGCTGCTGAGGCTGCACCT	4795
Db	5734	CATCACCACTCAACCAACCCCTTGAACCTTCAACAGGAGCAGCTGCTGAGGCTGCACCT	5793
Qy	4796	GATGGCCCTCTCGGTGGAGCTCTCGTGTCCATCTGTGTGGTCTTTTGGCCATGTCTTTGT	4855
Db	5794	GATGGCCCTCTCGGTGGAGCTCTCGTGTCCATCTGTGTGGTCTTTTGGCCATGTCTTTGT	5853
Qy	4856	CCGGCCAGCTTCACTCTTGTCTCTATTGAGGAGCAGTCAACCGAGCCAAAGCACCTGCA	4915
Db	5854	CCGGCCAGCTTCACTCTTGTCTCTATTGAGGAGCAGTCAACCGAGCCAAAGCACCTGCA	5913
Qy	4916	GCTCATGGGGGCGCTGTCCCGACCCCTTACTGTGGCTTGGCAATTTCTCTGGGACATGTG	4975
Db	5914	GCTCATGGGGGCGCTGTCCCGACCCCTTACTGTGGCTTGGCAATTTCTCTGGGACATGTG	5973
Qy	4976	TAACCTACTGTGCGACGATCGTGTGTCTCATCTTTCTGCGCTTCCAGCAGAGGGC	5035
Db	5974	TAACCTACTGTGCGACGATCGTGTGTCTCATCTTTCTGCGCTTCCAGCAGAGGGC	6033
Qy	5036	ATATGTGGCCCTGCCAACCTGCGTCTCTCTGCTGTGTGCTACTACTGTATGGCTGGTC	5095
Db	6034	ATATGTGGCCCTGCCAACCTGCGTCTCTCTGCTGTGTGCTACTACTGTATGGCTGGTC	6093
Qy	5096	GATCACACCGGTCAATGATCCAGCGCTCTTCTTCTTCTCGTGCGCAGCACAGCCTATGT	5155
Db	6094	GATCACACCGGTCAATGATCCAGCGCTCTTCTTCTTCTCGTGCGCAGCACAGCCTATGT	6153
Qy	5156	GGTGCTCACCTGCATAAACCTCTTTATTTGGCATCAATGGAAGATGGCCACCTTTGTGCT	5215
Db	6154	GGTGCTCACCTGCATAAACCTCTTTATTTGGCATCAATGGAAGATGGCCACCTTTGTGCT	6213
Qy	5216	TGAGCTCTTCTCTGATCAGAAAGTCGAGGAGGTGAGCCGGATCTTGAAACAGGTTCTCCT	5275
Db	6214	TGAGCTCTTCTCTGATCAGAAAGTCGAGGAGGTGAGCCGGATCTTGAAACAGGTTCTCCT	6273
Qy	5276	TATCTTCCCGACTTCTGTGCGCGGGGCTCAATTGAATGGTGGGAAACAGGCGCAT	5335
Db	6274	TATCTTCCCGACTTCTGTGCGCGGGGCTCAATTGAATGGTGGGAAACAGGCGCAT	6333
Qy	5336	GGCTGATGCTTTGAGCGCTTGGGACAGGAGTTCAGTCAACCTCGCGCTGGGAGGT	5395
Db	6334	GGCTGATGCTTTGAGCGCTTGGGACAGGAGTTCAGTCAACCTCGCGCTGGGAGGT	6393
Qy	5396	GGTCGGCAAGAACTCTTTGGCCATGGTGATACAGGGGCGCCCTCTTCTCTTCTTCACT	5455
Db	6394	GGTCGGCAAGAACTCTTTGGCCATGGTGATACAGGGGCGCCCTCTTCTCTTCTTCACT	6453
Qy	5456	ACTGCTGACACCGAAGCCAACTCTTGCCACAGCCGAGGTTAGGTTCTGTGCACTCTCT	5515
Db	6454	ACTGCTGACACCGAAGCCAACTCTTGCCACAGCCGAGGTTAGGTTCTGTGCACTCTCT	6513
Qy	5516	GGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGTTGGTCCAGGAGCCACCA	5575
Db	6514	GGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGTTGGTCCAGGAGCCACCA	6573
Qy	5576	GGGGGATGTGTGTGCTGCTAGGAACTTGACCAAGGTATACCGTGGGACAGGATGCCAGC	5635

Db 6574 GGGGATGTTGGTCTGAGAACTTGACCAAGGTATACCGTGGGACAGGATCCAGC 6633
QY 5636 TGTGACCGCTTGTGCTGGGATTTCCCTCGTGAAGTGTGTTGGCTGCTGGGTGTGAA 5695
Db 6634 TGTGACCGCTTGTGCTGGGATTTCCCTCGTGAAGTGTGTTGGCTGCTGGGTGTGAA 6693
QY 5696 TGGAGCAGGGAAGACGTCACGTTTTCGATGTGTGACGCGGGGACATTTGGCCAGCAGGG 5755
Db 6694 TGGAGCAGGGAAGACGTCACGTTTTCGATGTGTGACGCGGGGACATTTGGCCAGCAGGG 6753
QY 5756 CGAGCGTGTGCTGGGACGACGAGCTGCGGAGCCGAGGACCCAGTGTGCGCACCTCAGCAT 5815
Db 6754 CGAGCGTGTGCTGGGACGACGAGCTGCGGAGCCGAGGACCCAGTGTGCGCACCTCAGCAT 6813
QY 5816 GGGATACCTGCCCTCAATCGATGCCATCTTTGAGCTGTGACGCGGCCGAGCACCTCGA 5875
Db 6814 GGGATACCTGCCCTCAATCGATGCCATCTTTGAGCTGTGACGCGGCCGAGCACCTCGA 6873
QY 5876 GCTGCTTGGCGCTGCGCGGTGTCGCGAGGCCCGAGTGTGCCAGACGCTGGCTCAGG 5935
Db 6874 GCTGCTTGGCGCTGCGCGGTGTCGCGAGGCCCGAGTGTGCCAGACGCTGGCTCAGG 6933
QY 5936 CTTGCGCGCTGCGGACTCTCATGTGACGAGACCGGCTGCGAGCACCTACAGCGGAGG 5995
Db 6934 CTTGCGCGCTGCGGACTCTCATGTGACGAGACCGGCTGCGAGCACCTACAGCGGAGG 6993
QY 5996 GAAACAAACGCAAGCTGGCGGCGCTGCGCTGTTGGGACCCAGCGCTGTTTCT 6055
Db 6994 GAAACAAACGCAAGCTGGCGGCGCTGCGCTGTTGGGACCCAGCGCTGTTTCT 7053
QY 6056 GGACGAGCCGACCAAGGATGAGCCCGAGCGCGCGGCTTCTTTTGGAAACAGCTTTT 6115
Db 7054 GGACGAGCCGACCAAGGATGAGCCCGAGCGCGCGGCTTCTTTTGGAAACAGCTTTT 7113
QY 6116 GGCGTGTGCGGAGGCGCTTCACTGATGCTCATCTCCATAGCATGAGGAGTGGA 6175
Db 7114 GGCGTGTGCGGAGGCGCTTCACTGATGCTCATCTCCATAGCATGAGGAGTGGA 7173
QY 6176 AGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGGCGGTTCCGCTGCTGGGAGGCC 6235
Db 7174 AGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGGCGGTTCCGCTGCTGGGAGGCC 7233
QY 6236 CCAACATCTCAAGGSCAGATTGCGGCGGGTCAACACTGACCTTCCGCGGTGCGCGCGC 6295
Db 7234 CCAACATCTCAAGGSCAGATTGCGGCGGGTCAACACTGACCTTCCGCGGTGCGCGCGC 7293
QY 6296 AAGTCCAGCGGAGCGGCTTCTGTTGGCGGCGAGTTCCTTGGGTGCGAGCTGCGGA 6355
Db 7294 AAGTCCAGCGGAGCGGCTTCTGTTGGCGGCGAGTTCCTTGGGTGCGAGCTGCGGA 7353
QY 6356 GGACATGGAGGCGGCTTCCAGCTGCGCGGAGGCGGCTGCGGCGCTGGCGG 6415
Db 7354 GGACATGGAGGCGGCTTCCAGCTGCGCGGAGGCGGCTGCGGCGCTGGCGG 7413
QY 6416 CGTCTTTGGAGCTGCGGTCGACGCGGAGAGCAAGCGCTGGAGGACTTTTCCGTGAG 6475
Db 7414 CGTCTTTGGAGCTGCGGTCGACGCGGAGAGCAAGCGCTGGAGGACTTTTCCGTGAG 7473
QY 6476 CCAGACGATGCTGAGGAGGATTTCTTGTATCTTCTCAAGGACGAGGGAAGGACGAGGA 6535
Db 7474 CCAGACGATGCTGAGGAGGATTTCTTGTATCTTCTTCAAGGACGAGGGAAGGACGAGGA 7533
QY 6536 CACGAGAGCAGAGGAGGAGGAGTGGAGTGGAGTGGACCCCGCGCCGCGCTGAGCACCC 6595
Db 7534 CACGAGAGCAGAGGAGGAGGAGTGGAGTGGAGTGGACCCCGCGCCGCGCTGAGCACCC 7593
QY 6596 CAACCGCTCAGCCAGTTCCTCGATGACCTTAGCACTGCGGAGACTGTGCTCTGAGCCTC 6655
Db 7594 CAACCGCTCAGCCAGTTCCTCGATGACCTTAGCACTGCGGAGACTGTGCTCTGAGCCTC 7653
QY 6656 CCTCCCTGCGGCGCGCGGAGGCCCTTGGGAATGGCAAGGCAAGGTAGAGTGCCTAG 6715

Db 7654 CCTCCCTGCGGCGCGCGGAGGCCCTGGGAATGGCAAGGCAAGGTAGAGTGCCTAG 7713
QY 6716 GAGCCTCGAATCAGGCTGGCAGAGGGGCTGCTGCTGGCAAAATAAGAGAAAGGCTG 6775
Db 7714 GAGCCTCGAATCAGGCTGGCAGAGGGGCTGCTGCTGGCAAAATAAGAGAAAGGCTG 7773
QY 6776 GAGAGAGCCCTGCTGTGTGAAA 6797
Db 7774 GAGAGAGCCCTGCTGTGTGAAA 7795

RESULT 5
AF328787
LOCUS
DEFINITION
AF328787
ACCESSION
AF328787.1 GI:12656650
VERSION
AF328787.1
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eukarya; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 6704)
Broccardo, C., Osorio, J., Luciani, M.-F., Schriml, L.M., Prades, C.,
Shulenin, S., Arnould, I., Naudin, L., Lafargue, C., Rosier, M.,
Jordan, B., Mattei, M.G., Dean, M., Deneffe, P. and Chimini, G.
Comparative analysis of the promoter structure and genomic
organization of the human and mouse ABCA7 gene encoding a novel
ABCA transporter
Cytogenet. Cell Genet. 92 (3-4), 264-270 (2001)
MEDLINE
21328888
PUBMED
11435699
REFERENCE
2 (bases 1 to 6704)
Broccardo, C., Osorio, J., Luciani, M.-F., Lynn, S., Prades, C.,
Shulenin, S., Arnould, I., Naudin, L., Lafargue, C., Rosier, M.,
Jordan, B., Mattei, M.G., Dean, M., Deneffe, P. and Chimini, G.
Direct Submission
Submitted (13-DEC-2000) EGC, Aventis-Pharma, 2 Rue Gaston Cremieux,
EVRY, 91 91000, France
FEATURES
Location/Qualifiers
1..6704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1..6704
/gene="ABCA7"
/note="ABCA7 protein"
/codon_start=1
/product="ABC transporter member 7"
/protein_id="AAK00959.1"
/db_xref="GI:12656651"
/translation="MAFWTQMLLLWKNFMYRRQPVOLLVLLWLPFLPFLVAVRH
SHPLEHCHFPKPLPSAGTVPMLQGLICNVNNTCFPOLTGCEGRSLNFDLSV
SRLADARTVLGGASAHRTLAGLKLATLRAARSTAOPTKQSPLEPMLDVALL
TSLLRTESGLAQOEPHLSELAAGDLAQELLRSLVLRLLQRPRTGSPLE
LLEALCSVRGSPSTGSLNWEASDLMLVGQEPESALPDSLSFACSELIGALDS
HPLSLRLRLKPLILGLLPADPTPTKLMQVNRTEBELTLRDVREWELGPR
IPTFMNDSNVAMLORLLOMDGRRQPRPGDRHMEALRSFLDPGGSGYSWQADAH
VGHVCTGLVTECLSLDLKLEAASEALVSRALQLAEHRFWAGVFLGPESSDPT
EHTPDLGPHVRIKRMIDVTRINKIRDFWDPGPAADPLTDLRYWGGFFVLQD
LVBERAVRLUSGANPRAGLYLQMPYPCYVDVFLVLSRLSLPLFLTLAMIYSLTIV
KAVVREKTERLDRMTMRAGLSRAVLWGLFSLCLGPPFLSALALVLFLTLAMIYS
PGVVFLLAFAVATVTSPELLSAFFSRANLAAACGLAYFSLVPLVLCVAMDRLLP
AGRVAASLLSPVAFGFCESLALLEEQGGAQHNVTGPTADVLSLAQVSGILLD
AALYGLATVLEAVCPQGYGIPFPNPPFRSFWCGRPKSPACPTPLDPKVLVBE
APPLSPGVSRSLERKFPQSPALRGLSLDFQGHITAPLGHNGAGKTTLSILSG
LPPGSGSAPFILGHVRSMAAIPHLGVCPOYNVLFMDLTVDEHVMFWYGRLLSLAA
VVGPEQDRLIQDVLGYSKQSVQTRHLSGGMQRKLSVAIAFVGGSQVVLDEPTAGVDP
ASRRGIWELLKYRBEGRTLTSLTHHLD EALLGLDRVAVAGGRLCCCGSPLFLRRHLG

SGYLLTLVKARLPLTTNEKADTDMEGSDVTRQEKNGKSGSRVGTGPOLLALVOHWVPG
ARLVEELPHELVLVLPVTHAGDGSFATLRELRLAELRLTCVGTSDTSLEELPLKV
VEECADTDWEDSGSCQHLLCTGAGLDVTLRLKMPQETALENGEPAGNAPETDQSG
PDVAVRQVGHALLTQQLALLKRFLLARRSRRLGFAQIVLPALFVGLALVFLSLI VPP
FGHPALRUSPTMYGAQVSPFSDDA PGDPRARLLLEALLQEALEGPVVOHSSHRFSA
PEYPAEYKVLASGNWTPESPSPACQSRPGARRLLPDCPAAAGPFPQPAVTGSGEV
VQNLTRNLSDFLVKVTPRLVROGLTKKWNVRYGFSLGGDRDPLSGQELGRSV
EELWALLSPPLGGALORVNLKNTAWAHSIDAQDSLAKI WFNKNGHWSMVA FVNRASNAI
LRAHLPGPAPARHAHSITTLNHPNLNKEQLSEAAALMASSVDVLVSICVVPANSPVPAS
FTVLVEERVIRAHGLQMGGLSFTLWGLNFWDMCNLYVPACI WLLFLAPQQRAY
VAPANILPALLLLLLGWSITPLMYPASFSPFSVSTAYVVLTCINLIFIGINSMAITFV
LELFSDQKLEVERIILKQVFLI FPHFCLGRGLDMVRNQAMADAFERLGDORQPSRLR
WEVGNKLLAMVQGLPFLI LFTLLQHSOLLPPQVRSLPILLEGEDEDVARERVRV
QAGTQDVLVLRNLTKVYRGORMPADVRLCLGIPGBCFCGLLGVNGCAKSTPRMTYG
DTLASGEAVLASHVAREPSAAHLSMGYCPQSDAIPELLTGREHLELLARLGRVPEA
QVAQTGASGLRGLSWYADRPAGTYSGGNKRKLATALALVDPDAVVFIDEPTGMDP
SARFLWNSLLAVRGRSVMYLTSHSMECEBALCSRLAIVMNGRFRCLGSPQHLGRF
AAGHTTLTRVPAARSOPAAAFVAEPFSGELREAHGRRLFPQPPGRCALARVFGEL
AVHGABHGVDFSVSOTMLEBEVFLYFSKQGDDETEBEQKEAGVGVDPAPGLQHPKRV
SQFLDDPSTAEITVL"

ORIGIN

Query Match 98.3%; Score 6686.4; DB 9; Length 6704;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6693; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy	94	GTGGGATAAAGGAATGAGGTTTCAGAAAGGGCGAGGAGTTGCCGCGACGCCGACCGGACG 153
Db	1	GTGGGATAAAGGAATGAGGTTTCAGAAAGGGCGAGGAGTTGCCGCGACGCCGACCGGACG 60
Qy	154	TCCTTACGCCGACCGTGTCTGACACCTCTGTCCGTCGCCGCCGCCGACGTCACCATGG 213
Db	61	TCCTTACGCCGACCGTGTCTGACACCTCTGTCCGTCGCCGCCGCCGACGTCACCATGG 120
Qy	214	CTTCTTGGACA CAGCTGATGCTGCTGCTCTGGAAGAAATTTTCATGTATCGCCGAGACAGC 273
Db	121	CTTCTTGGACA CAGCTGATGCTGCTGCTCTGGAAGAAATTTTCATGTATCGCCGAGACAGC 180
Qy	274	CGGTCCAGCTCTGCTGTAATTTGCTGTGGCTCTCTTCTCTCTTCTTCTCTCTCTCTCTCT 333
Db	181	CGGTCCAGCTCTGCTGTAATTTGCTGTGGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy	334	TTGCCACTCCACCCGCCCTTGAGCACCATGAA TGCCACTTCCCAAAACAGCACTGC 393
Db	241	TTGCCACTCCACCCGCCCTTGAGCACCATGAA TGCCACTTCCCAAAACAGCACTGC 300
Qy	394	CATCGGCGGACCGTGTCCAGGCTCTCATCTGTAAATGTGAACAA CACCTGCT 453
Db	301	CATCGGCGGACCGTGTCCAGGCTCTCATCTGTAAATGTGAACAA CACCTGCT 360
Qy	454	TTCCGAGCTGACACCGGCGAGGACCGCGGCGCTTGACAACTTCAACGACTCCCTTG 513
Db	361	TTCCGAGCTGACACCGGCGAGGACCGCGGCGCTTGACAACTTCAACGACTCCCTTG 420
Qy	514	TCCTCCGCTGTAGCGGATCCCGCACTGTCTGTGGAGGGGCGAGTGCCCA CAGGACGC 573
Db	421	TCCTCCGCTGTAGCGGATCCCGCACTGTCTGTGGAGGGGCGAGTGCCCA CAGGACGC 480
Qy	574	TGGCTGCGCTAGGGAAGCTGATCGCCACGCTGAGGGCTGCA CGCAGCA CGGCCCGAGCCTC 633
Db	481	TGGCTGCGCTAGGGAAGCTGATCGCCACGCTGAGGGCTGCA CGCAGCA CGGCCCGAGCCTC 540
Qy	634	AACCAACCAAGCTTCCA CTGGAACCA CCAATGCTGGAATGTGCGGAGCTGTGACGT 693
Db	541	AACCAACCAAGCTTCCA CTGGAACCA CCAATGCTGGAATGTGCGGAGCTGTGACGT 600
Qy	694	CACCTGCTGCGACGGAATCCCTGGGTTGCACTTGGSCCAAGCCACGAGSCCTTTGCACA 753
Db	601	CACCTGCTGCGACGGAATCCCTGGGTTGCACTTGGSCCAAGCCACGAGSCCTTTGCACA 660
Qy	754	GCTTGTGGAGGCGCTGAGGACCTGGGCCAGGAGCTCTCTGGGCTGCGGAGCTGTGTTGG 813
Db	661	GCTTGTGGAGGCGCTGAGGACCTGGGCCAGGAGCTCTCTGGGCTGCGGAGCTGTGTTGG 720

Qy	814	AGCTTCGGGCACTGCTGCAGAGACCCCGAGGAGACAGCGCGCCCTTGAGTTGCTGTCTCAG 873
Db	721	AGCTTCGGGCACTGCTGCAGAGACCCCGAGGAGACAGCGCGCCCTTGAGTTGCTGTCTCAG 780
Qy	874	AGSCCCTCTGCAGTGTTCAGGGGACCTAGCAGACACAGTGGGCCCCCTCAACTGTGTACG 933
Db	781	AGSCCCTCTGCAGTGTTCAGGGGACCTAGCAGACACAGTGGGCCCCCTCAACTGTGTACG 840
Qy	934	AGCTTAGTACCTGATGGAGCTGTGGGGCAGGAGCCAGAA TCCGCCCTTCCGACAGCA 993
Db	841	AGCTTAGTACCTGATGGAGCTGTGGGGCAGGAGCCAGAA TCCGCCCTTCCGACAGCA 900
Qy	994	GCCTGAGCCCCGCTCTCTCGAGCTCATTTGGAGCCCTTGACAGCCACCGCTCTTCCGCC 1053
Db	901	GCCTGAGCCCCGCTCTCTCGAGCTCATTTGGAGCCCTTGACAGCCACCGCTCTTCCGCC 960
Qy	1054	TGCTCTGGAGACGCGCTTGAAGCCCTCTGATCCTCGGGAAGCTACTCTTTTGCA CCA GATACAC 1113
Db	961	TGCTCTGGAGACGCGCTTGAAGCCCTCTGATCCTCGGGAAGCTACTCTTTTGCA CCA GATACAC 1020
Qy	1114	CTTTTACC CGGAAGCTCATGSCCCAGGTGAA CCGGACCTTTCGAGGAGCTCACCCCTGCTGA 1173
Db	1021	CTTTTACC CGGAAGCTCATGSCCCAGGTGAA CCGGACCTTTCGAGGAGCTCACCCCTGCTGA 1080
Qy	1174	GGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTTCATCTCATGAA CCGACA 1233
Db	1081	GGGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTTCATCTCATGAA CCGACA 1140
Qy	1234	GTTCCTCAATGTGSCCATGTCTGCAGCGCTCTCTGCAGATGAGGATGAAGAAAGGAGCAGC 1293
Db	1141	GTTCCTCAATGTGSCCATGTCTGCAGCGCTCTCTGCAGATGAGGATGAAGAAAGGAGCAGC 1200
Qy	1294	CCAGACTCTGGAGGCGGGGACCA CATGGAGGCGCTCTGCGATCTCTTTCTGAGACCTCTGGAGCG 1353
Db	1201	CCAGACTCTGGAGGCGGGGACCA CATGGAGGCGCTCTGCGATCTCTTTCTGAGACCTCTGGAGCG 1260
Qy	1354	GTGGCTACAGCTGGGAGGACGCA CACGCTGATGTGGGGCACTGTGTGGGCGCGCTCTGGG 1413
Db	1261	GTGGCTACAGCTGGGAGGACGCA CACGCTGATGTGGGGCACTGTGTGGGCACTGTGTGGGCG 1320
Qy	1414	GAGTGA CCGAGTGCTCTGCTTCTTGGCA CAAGCTGGAGGCGGCA CCGCTCAGAGGCGACCTGG 1473
Db	1321	GAGTGA CCGAGTGCTCTGCTTCTTGGCA CAAGCTGGAGGCGGCA CCGCTCAGAGGCGACCTGG 1380
Qy	1474	TGTTCGGGCGCTGTGCAATGTCTCGCGGAA CACATCGATTCTGGGCGGCGCTGTCTTCTTGG 1533
Db	1381	TGTTCGGGCGCTGTGCAATGTCTCGCGGAA CACATCGATTCTGGGCGGCGCTGTCTTCTTGG 1440
Qy	1534	GACCTGAGGACTCTTTCAGACCCCA CAGACGACCCCAACCCAGACCTGGGCGCGGCGCACG 1593
Db	1441	GACCTGAGGACTCTTTCAGACCCCA CAGACGACCCCAACCCAGACCTGGGCGCGGCGCACG 1500
Qy	1594	TGCGCATCAAAATCCGCATGGA CATTTGACGTGTGTCA CAGAGGACCAATAAGATCAGGAGCA 1653
Db	1501	TGCGCATCAAAATCCGCATGGA CATTTGACGTGTGTCA CAGAGGACCAATAAGATCAGGAGCA 1560
Qy	1654	GGTTTGGGACCCCTGCGCAGCGCGGAC CCGCTCTGA CCGACCTGCGCTACGTGTGGGGCG 1713
Db	1561	GGTTTGGGACCCCTGCGCAGCGCGGAC CCGCTCTGA CCGACCTGCGCTACGTGTGGGGCG 1620
Qy	1714	GCTTCGTGTACCTGCAAGACCTGTGGAGGCTGACGCGCTCCGCGTCTCAGCGGCGCCA 1773
Db	1621	GCTTCGTGTACCTGCAAGACCTGTGGAGGCTGACGCGCTCCGCGTCTCAGCGGCGCCA 1680
Qy	1774	ACCCCGGGCGCGCTCTACCTGCAGCAGATGCCCTTATCCGTGCTATGTGGA CCA CAGTGT 1833
Db	1681	ACCCCGGGCGCGCTCTACCTGCAGCAGATGCCCTTATCCGTGCTATGTGGA CCA CAGTGT 1740
Qy	1834	TCTTCGTGTGTGACCGGCTGCTGCGCTCTTCTCTGACGCTGTGCGCTGTGATCTACTCCG 1893
Db	1741	TCTTCGTGTGTGACCGGCTGCTGCGCTCTTCTCTGACGCTGTGCGCTGTGATCTACTCCG 1800

Qy	1894	TGACACTGACAGTGAAGGCGGTGCTGGGGAGAGAGGAGACGGCGTGGGGGACACCAATGC	1954
Db	1801	TGACACTGACAGTGAAGGCGGTGCTGGGGAGAGAGGAGACGGCGTGGGGGACCAATGC	1860
Qy	1954	GCGCCATGGGGCTCAGCGCGCGGTGCTTGGCTAGGCTGGTTCCTCAGCTGCCTCGGGC	2013
Db	1861	GCGCCATGGGGCTCAGCGCGCGGTGCTTGGCTAGGCTGGTTCCTCAGCTGCCTCGGGC	1920
Qy	2014	CCTTCCTGCTCAGCGCGCGGTGCTGGTTCTGGTGTCTAAGCTGGGGGACATCTCCCTC	2073
Db	1921	CCTTCCTGCTCAGCGCGCGCATGCTGGTTCTGGTGTCTAAGCTGGGAGACATCTCCCTC	1980
Qy	2074	ACAGCCACCCGGGGGTGCTTCTGCTTCTGGCAGAGCTTTCGGGTGGCCACCGTGAACC	2133
Db	1981	ACAGCCACCCGGGGGTGCTTCTGCTTCTGGCAGAGCTTTCGGGTGGCCACCGTGAACC	2040
Qy	2134	AGAGCTTCCTGCTCAGCGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCTGGGGCGGCC	2193
Db	2041	AGAGCTTCCTGCTCAGCGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCTGGGGCGGCC	2100
Qy	2194	TGGCCTACTTCTCCCTCTACCTGCCCTACGTGTCTGTGTGGCTTGGCGGACCGGTGC	2253
Db	2101	TGGCCTACTTCTCCCTCTACCTGCCCTACGTGTCTGTGTGGCTTGGCGGACCGGTGC	2160
Qy	2254	CCGGGGTGGCGGCTGGCGCGAGCCTGTGTGTCCCGGTGGCCCTTGGGCTTGGGCTGCG	2313
Db	2161	CCGGGGTGGCGGCTGGCGCGAGCCTGTGTGTCCCGGTGGCCCTTGGGCTTGGGCTGCG	2220
Qy	2314	AGAGCTTGGCTTCTCTGAGGAGCAGGGCGAGGGCGCGAGTGGGCAACAGTGGGGACCC	2373
Db	2221	AGAGCTTGGCTTCTCTGAGGAGCAGGGCGAGGGCGCGAGTGGGCAACAGTGGGGACCC	2280
Qy	2374	GGCCTACGGCAGAGCTTTACGCTGSCCCAGAGTCTCTGSCCTTCTGCTCTG3ACGCGG	2433
Db	2281	GGCCTACGGCAGAGCTTTACGCTGSCCCAGAGTCTCTGSCCTTCTGCTCTG3ACGCGG	2340
Qy	2434	CGCTACGGGCTCGCACCTGTGTACCTGGAAGCTGTGTGCCAGGCGAGTACGGGATCC	2493
Db	2341	CGCTACGGGCTCGCACCTGTGTACCTGGAAGCTGTGTGCCAGGCGAGTACGGGATCC	2400
Qy	2494	CTGAACCATGGAAATTTTCTTTTCGGAGGAGCTACTGCTGGGACCTCGGCCCCCAAGA	2553
Db	2401	CTGAACCATGGAAATTTTCTTTTCGGAGGAGCTACTGCTGGGAGCTCGGCCCCCAAGA	2460
Qy	2554	GTCAGAGCCCTTGCCCAACCCGCTGGACCCAAAGGTGCTGTAGAGAGGCAACCGCCG	2613
Db	2461	GTCAGAGCCCTTGCCCAACCCGCTGGACCCAAAGGTGCTGTAGAGAGGCAACCGCCG	2520
Qy	2614	GCCTGAGTCTGGGCTATCCGTCGACCTTGAGAAAGCGCTTTCGTGAAGGCCCGCAGC	2673
Db	2521	GCCTGAGTCTGGGCTATCCGTCGACCTTGAGAAAGCGCTTTCGTGAAGGCCCGCAGC	2580
Qy	2674	CAGCCCTGCGGGGCTCAGCTGGACTTCTACAGGGCGCATACCGCCTTCTTGGGCC	2733
Db	2581	CAGCCCTGCGGGGCTCAGCTGGACTTCTACAGGGCGCATACCGCCTTCTTGGGCC	2640
Qy	2734	ACAACGGGGCGGCAAGACCAACACCTGTCCATCTTGAAGTGGCCCTTTCACACCCAGTG	2793
Db	2641	ACAACGGGGCGGCAAGACCAACACCTGTCCATCTTGAAGTGGCCCTTTCACACCCAGTG	2700
Qy	2794	GTGGCTCTGCTTATCTCGGGGCAACGCTGCGCTCCAGCATGGCGCCCATCCGGCCCC	2853
Db	2701	GTGGCTCTGCTTATCTCGGGGCAACGACGCTCCGCTCCAGCATGGCGCCCATCCGGCCCC	2760
Qy	2854	ACCTGGGGCTGTGCTTCAGTACAAAGTGTGTTTGACATGCTGAACGTTGAAGAGCAG	2913
Db	2761	ACCTGGGGCTGTGCTTCAGTACAAAGTGTGTTTGACATGCTGAACGTTGAAGAGCAG	2820
Qy	2914	TCTGGTTCTATGGCGGCTGAAGGGTCTGAGTGGCGCTGTAGTGGGCGCCGAGCAGGACC	2973
Db	2821	TCTGGTTCTATGGCGGCTGAAGGGTCTGAGTGGCGCTGTAGTGGGCGCCGAGCAGGACC	2880
Qy	2974	GTCTGCTCAGGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGACTCGCACCTCTCTG	3033

Db	2881		GTCTGCTGCAGGATGTGGGGCTGGTCTCAAGACAGAGTGCAGACTCGCACCTCTCTG	2940
Qy	3034		GTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCTCTTTGTGGGGCGCTCCCAAGTTGTTA	3093
Db	2941		GTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCTCTTTGTGGGGCGCTCCCAAGTTGTTA	3000
Qy	3094		TCTGTGACGAGCTTACGGCTGCGTGTGATCTCTGTCTTCCGCGCGCGTATTTGGGAGCTGC	3153
Db	3001		TCTGTGACGAGCTTACGGCTGCGTGTGATCTCTGTCTTCCGCGCGCGTATTTGGGAGCTGC	3060
Qy	3154		TGCTCAAAATACCGAAGGTGCACGCTGATCTCTCCACCCACACCTGATGAGGCAG	3213
Db	3061		TGCTCAAAATACCGAAGGTGCACGCTGATCTCTCCACCCACACCTGATGAGGCAG	3120
Qy	3214		AGTGTCTGGGAGACCGTGTGGCTGTGTGGGAGGTGGCGGCTTGTGCTGTGCTGGCTCCC	3273
Db	3121		AGTGTCTGGGAGACCGTGTGGCGGTGTGGGAGGTGGCGGCTTGTGCTGTGCTGGCTCCC	3180
Qy	3274		CACCTTCTCTGCGCGTCACTTGGGCTCCGGCTACTACTGACGCTGTGTGAAGCCCGCC	3333
Db	3181		CACCTTCTCTGCGCGTCACTTGGGCTCCGGCTACTACTGACGCTGTGTGAAGCCCGCC	3240
Qy	3334		TGCCCCGTGACCAATGAGAGGCTGACATGACATGGAGGGCAGTGTGACACACGAGC	3393
Db	3241		TGCCCCGTGACCAATGAGAGGCTGACATGACATGGAGGGCAGTGTGACACACGAGC	3300
Qy	3394		AGGAAAGAAGAAATGGCAGCCAGGCGACAGAGTCCGGCACTCTCTCAGCTGTGCTGCCCTGG	3453
Db	3301		AGGAAAGAAGAAATGGCAGCCAGGCGACAGAGTCCGGCACTCTCTCAGCTGTGCTGCCCTGG	3360
Qy	3454		TACAGCACTGGGTGTCGCCGGGCAACCGCTGGTGGAGGAGCTGCCACACGAGCTGTGCTGG	3513
Db	3361		TACAGCACTGGGTGTCGCCGGGCAACCGCTGGTGGAGGAGCTGCCACACGAGCTGTGCTGG	3420
Qy	3514		TGCTGCCCTACACGGGTGCCCATGACGCGAGCTTGCACACACTCTTCCGAGAGCTAGACA	3573
Db	3421		TGCTGCCCTACACGGGTGCCCATGACGCGAGCTTGCACACACTCTTCCGAGAGCTAGACA	3480
Qy	3574		CGCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCCTCGAGAGAGA	3633
Db	3481		CGCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCCTCGAGAGAGA	3540
Qy	3634		TCTTCTGTAAGTGGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGCTGCG	3693
Db	3541		TCTTCTGTAAGTGGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGCTGCG	3600
Qy	3694		GGCAGCACTATGCAAGGCAATGTGCGCTTAGAGCTAAACCTGGGGCTCAAGATGCCGC	3753
Db	3601		GGCAGCACTATGCAAGGCAATGTGCGCTTAGAGCTAAACCTGGGGCTCAAGATGCCGC	3660
Qy	3754		CACAGAGACGCGTGGAGAACGGGGAACAGCTGGGTACGCCCCAGAGACTCACGAGG	3813
Db	3661		CACAGAGACGCGTGGAGAACGGGGAACAGCTGGGTACGCCCCAGAGACTCACGAGG	3720
Qy	3814		GCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGGGGCACTGACCCGCGCAGAGTCC	3873
Db	3721		GCTCTGGGCCAGACGCGTGGGCCGGGTACAGGGCTGGGGCACTGACCCGCGCAGAGTCC	3780
Qy	3874		AGGCCCTGTCTTCAAGGCGCTTTCTGCTTGGCCGCGCAGCGCGCGGCTGTGTTGCGCC	3933
Db	3781		AGGCCCTGTCTTCAAGGCGCTTTCTGCTTGGCCGCGCAGCGCGCGGCTGTGTTGCGCC	3840
Qy	3934		AGATCGTGTGCTCGCCCTCTTTGTGGGCCCTGGGCCCTCGTGTTCAGGCTCATCTCGTGCTC	3993
Db	3841		AGATCGTGTGCTCGCCCTCTTTGTGGGCCCTGGGCCCTCGTGTTCAGGCTCATCTCGTGCTC	3900
Qy	3994		CTTTTGGGCACTACCCGGCTCTGCGGCTCAGTCCACCATTACGGTGTCTCAGGTGTCT	4053
Db	3901		CTTTTGGGCACTACCCGGCTCTGCGGCTCAGTCCACCATTACGGTGTCTCAGGTGTCT	3960
Qy	4054		TCTTCAGTGAGGACGCCCCAGGGGACCTTGACGTGCGCGCTGCTCGAGGCGCTGTGTC	4113

Db	3961	TCCTCAGTGAGGACGCCCCAGGGACCCCTGAGCGTCCCGCGCTGCTCGAGCGCTGCTGC	4020
Qy	4114	AGGAGGAGGACTGGAGAGCCCCCAGTGCAGCATAGCTCCCAAGGTTCTCGGCACAG	4173
Db	4021	AGGAGGAGGACTGGAGAGCCCCCAGTGCAGCATAGCTCCCAAGGTTCTCGGCACAG	4080
Qy	4174	AAGTTCTCTGTGAAGTGGCCAAAGGTCTTGGCCAGTGCAACTGGACCCCAAGAGTCTCCAT	4233
Db	4081	AAGTTCTCTGTGAAGTGGCCAAAGGTCTTGGCCAGTGCAACTGGACCCCAAGAGTCTCCAT	4140
Qy	4234	CCCAGCGCTGCAGTGTAGCCAGCCCGGTGCCCGGCCCTGCTGCCGAGCTGCCCGGCTG	4293
Db	4141	CCCAGCGCTGCAGTGTAGCCAGCCCGGTGCCCGGCCCTGCTGCCGAGCTGCCCGGCTG	4200
Qy	4294	CAGCTGTGTGTCCTCCCTCCGCCCCAGGCAGTGCAGCGCTCTGGGGAAGTGGTTCAGAAC	4353
Db	4201	CAGCTGTGTGTCCTCCCTCCGCCCCAGGCAGTGCAGCGCTCTGGGGAAGTGGTTCAGAAC	4260
Qy	4354	TGACAGGCCCGAACTCTGTCTGACTTTCCTGTGTCAGAGCTTACCCCGCGCTGCTGCCAGG	4413
Db	4261	TGACAGGCCCGAACTCTGTCTGACTTTCCTGTGTCAGAGCTTACCCCGCGCTGCTGCCAGG	4320
Qy	4414	GCCTGAAGACTAAGAACTGGGTGAATGAGTTCAGGTACGAGGCTTCTCCTCGTGGGGGCC	4473
Db	4321	GCCTGAAGACTAAGAACTGGGTGAATGAGTTCAGGTACGAGGCTTCTCCTCGTGGGGGCC	4380
Qy	4474	GAGACCCAGGCTGCCCTCCGCCCAAGATTTGGCCCGCTCAGTGGAGAGTGTGGGCGC	4533
Db	4381	GAGACCCAGGCTGCCCTCCGCCCAAGATTTGGCCCGCTCAGTGGAGAGTGTGGGCGC	4440
Qy	4534	TGCTAGTTCCTCCCTGGGGGGCCCTCGACCGTGTCTTGAAGAACTCAGACGCTGGG	4593
Db	4441	TGCTAGTTCCTCCCTGGGGGGCCCTCGACCGTGTCTTGAAGAACTCAGACGCTGGG	4500
Qy	4594	CTCAGCCTGGATGTCTCAGSACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACT	4653
Db	4501	CTCAGCCTGGATGTCTCAGSACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACT	4560
Qy	4654	CCATGTGGCTTTGTCAAACGAGCCAGCAACGCAATCTCTCGTGTCTACCTGCCCCCAG	4713
Db	4561	CCATGTGGCTTTGTCAAACGAGCCAGCAACGCAATCTCTCGTGTCTACCTGCCCCCAG	4620
Qy	4714	GCCCGGCCCGCCAGCCCAAGCATACCACTCAACACCCCTTGAACCTCACCAGG	4773
Db	4621	GCCCGGCCCGCCAGCCCAAGCATACCACTCAACACCCCTTGAACCTCACCAGG	4680
Qy	4774	AGCAGCTGTCTGAGGCTGCACATGATGGCCCTCTCGGTGGACGTCCTCGTCTCCAATCTGTG	4833
Db	4681	AGCAGCTGTCTGAGGCTGCACATGATGGCCCTCTCGGTGGACGTCCTCGTCTCCAATCTGTG	4740
Qy	4834	TGGTCTTTGGCATGTCTTTGTCTCCGGCCAGCTTCACTCTGTGCTCAATTTAGAGGCGAG	4893
Db	4741	TGGTCTTTGGCATGTCTTTGTCTCCGGCCAGCTTCACTCTGTGCTCAATTTAGAGGCGAG	4800
Qy	4894	TCACCCGAGCCAGACCTGCAGCTCATGGGGGCCCTGTCCTCCACCCCTCTACTGGCTTG	4953
Db	4801	TCACCCGAGCCAGACCTGCAGCTCATGGGGGCCCTGTCCTCCACCCCTCTACTGGCTTG	4860
Qy	4954	GCAACTTTCTCTGGGACATGTGAACCTACTTGTGTGCCAGCATGTCATGCTGTCTCATCT	5013
Db	4861	GCAACTTTCTCTGGGACATGTGAACCTACTTGTGTGCCAGCATGTCATGCTGTCTCATCT	4920
Qy	5014	TTCTGGCTTTCCAGCAGAGGCGCATATGTGGCCCTTGCCAACTGCTGCTCTCTGCTGT	5073
Db	4921	TTCTGGCTTTCCAGCAGAGGCGCATATGTGGCCCTTGCCAACTGCTGCTCTCTGCTGT	4980
Qy	5074	TGCTACTACTGTATGGCTGTGATACACCGCTCATGTACCCAGCCCTCTCTCTTCT	5133
Db	4981	TGCTACTACTGTATGGCTGTGATACACCGCTCATGTACCCAGCCCTCTCTCTTCT	5040
Qy	5134	CCGTGCCAGCAGCCCTATGTGTGCTCAGCTGCATAAACCTCTTTTATTTGGCATCAATG	5193
Db	5041	CCGTGCCAGCAGCCCTATGTGTGCTCAGCTGCATAAACCTCTTTTATTTGGCATCAATG	5100

Qy	5194	GAAGCATGGCCACCTTTGTGCTTGTAGCTCTTCTCTGATCAGAAAGCTGCAGAGGTGAGCC	5253
Db	5101	GAAGCATGGCCACCTTTGTGCTTGTAGCTCTTCTCTGATCAGAAAGCTGCAGAGGTGAGCC	5160
Qy	5254	GGATCTTGAACCAAGGCTTCTCTTATCTTCCCCCACTTCTCTTGGGCGGGGCTCATTTG	5313
Db	5161	GGATCTTGAACCAAGGCTTCTCTTATCTTCCCCCACTTCTCTTGGGCGGGGCTCATTTG	5220
Qy	5314	ACATGTGCGGAAACCAAGGCGCATGGCTGATGCTTTTGTAGCCCTTTGGGAGACAGGAGTTCC	5373
Db	5221	ACATGTGCGGAAACCAAGGCGCATGGCTGATGCTTTTGTAGCCCTTTGGGAGACAGGAGTTCC	5280
Qy	5374	AGTCAACCCCTGCGCTGGGAGTGGTGGCAAGAACTCTTGGCCATGTTGATACAGGGC	5433
Db	5281	AGTCAACCCCTGCGCTGGGAGTGGTGGCAAGAACTCTTGGCCATGTTGATACAGGGC	5340
Qy	5434	CCCTCTTCTTCTTCTTCACTACTGCTGAGCAGCAGAAAGCCAACTCTGCCACAGCCCA	5493
Db	5341	CCCTCTTCTTCTTCTTCACTACTGCTGAGCAGCAGAAAGCCAACTCTGCCACAGCCCA	5400
Qy	5494	GGTGAGGCTCTGTGCCACTCTCTGGGAGAGGAGCAGAGGATGTAGCCCTGTGAAACGGGAGC	5553
Db	5401	GGTGAGGCTCTGTGCCACTCTCTGGGAGAGGAGCAGAGGATGTAGCCCTGTGAAACGGGAGC	5460
Qy	5554	GGGTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTGTCTGAGGAACTTGACCAAGTAT	5613
Db	5461	GGGTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTGTCTGAGGAACTTGACCAAGTAT	5520
Qy	5614	ACCGTGGGAGAGGATGCCAGCTGTTCACCGCTTGTGCTTGGGGATTTCCCTCTGGTGTAGT	5673
Db	5521	ACCGTGGGAGAGGATGCCAGCTGTTCACCGCTTGTGCTTGGGGATTTCCCTCTGGTGTAGT	5580
Qy	5674	GTTTTGGGCTGTGGTGTGAATGGAGCAGGAGGAAAGCTCCAAGTTTCGATGGTGTACGG	5733
Db	5581	GTTTTGGGCTGTGGTGTGAATGGAGCAGGAGGAAAGCTCCAAGTTTCGATGGTGTACGG	5640
Qy	5734	GGACACATTTGGCCAGCGGGGGAGGCTGTGCTGCGAGCCACAGGTCGGCCCGGGAAC	5793
Db	5641	GGACACATTTGGCCAGCGGGGGAGGCTGTGCTGCGAGCCACAGGTCGGCCCGGGAAC	5700
Qy	5794	CCAGTGTGCGCACCTCAGCATGGGATATCTGCCCTCAAATCCGATGCCATCTTTTGTAGCTGC	5853
Db	5701	CCAGTGTGCGCACCTCAGCATGGGATATCTGCCCTCAAATCCGATGCCATCTTTTGTAGCTGC	5760
Qy	5854	TGACGGGCGCGAGACCTGTGAGCTGCTGTGCGCGCTGTGCGCGTGTCCCGAGGCGCCAGG	5913
Db	5761	TGACGGGCGCGAGACCTGTGAGCTGCTGTGCGCGCTGTGCGCGTGTCCCGAGGCGCCAGG	5820
Qy	5914	TTGCCCCAGACCGCTGGCTCAGGCGCTCGCGCTCTGGGACTCTCATGGTACGAGACCGGC	5973
Db	5821	TTGCCCCAGACCGCTGGCTCAGGCGCTCGCGCTCTGGGACTCTCATGGTACGAGACCGGC	5880
Qy	5974	CTGACGACACTTACGCGGAGGAAACAAAGCTGGCGAGCGGCCCTGGCGCTGGTTG	6033
Db	5881	CTGACGACACTTACGCGGAGGAAACAAAGCTGGCGAGCTGGCGAGCGGCCCTGGCGCTGGTTG	5940
Qy	6034	GGACCCAGCGGTGTGTTTCTGGACGAGCGAGACCAAGCATGGAGCCCGAGCGCGCGC	6093
Db	5941	GGACCCAGCGGTGTGTTTCTGGACGAGCGAGACCAAGCATGGAGCCCGAGCGCGCGC	6000
Qy	6094	GCTTCTTTGGACACGCTTTTGGCGTGGTGGGAGGGCCCTCAGTGTGATGCTCACCT	6153
Db	6001	GCTTCTTTGGACACGCTTTTGGCGTGGTGGGAGGGCCCTCAGTGTGATGCTCACCT	6060
Qy	6154	CCCATAGCATGGAGGAGTGTGAAGCGCTCTGTCTGGCGCTTAGCCATCATGGTGAATGGGC	6213
Db	6061	CCCATAGCATGGAGGAGTGTGAAGCGCTCTGTCTGGCGCTTAGCCATCATGGTGAATGGGC	6120
Qy	6214	GGTTCCGCTGCTGGGAGCGCCCGAACTCATTAAGGGCAGATTCGCGGGGGTGCACAC	6273
Db	6121	GGTTCCGCTGCTGGGAGCGCCCGAACTCATTAAGGGCAGATTCGCGGGGGTGCACAC	6180

Qy	2790	AGTGGTGGCTCTGCTTCACTCTGGGCCACGACGTTCGCTCCAGCATGGCGCCCATCCGG	2849
Db	2581	AGTGGTGGCTCTGCTTCACTCTGGGCCACGACGTTCGCTCCAGCATGGCGCCCATCCGG	2640
Qy	2850	CCCACCTTGGCGCTGCTCCTCAGTACAAACGTGCTGTTTGACATGCTGACCGGTGGACGAG	2909
Db	2641	CCCACCTTGGCGCTGCTGCTCAGTACAAACGTGCTGTTTGACATGCTGACCGGTGGACGAG	2700
Qy	2910	CACGTCGTGTTCTATATGGGCGGCTGAAGGGTCTGAGTGC CGCTGTAGTGGGCGCCGAGCAG	2969
Db	2701	CACGTCGTGTTCTATATGGGCGGCTGAAGGGTCTGAGTGC CGCTGTAGTGGGCGCCGAGCAG	2760
Qy	2970	GACCGTCTGTCAGGATATGGGCGCTGTTCTCAAGCAGAGTGTGACACTCGCCACCTC	3029
Db	2761	GACCGTCTGTCAGGATATGGGCGCTGTTCTCAAGCAGAGTGTGACACTCGCCACCTC	2820
Qy	3030	TCTGTTGGGATGCAACGGAGCTGTCGTGGGCCATTCCTTTTGTGGGCGGCTCCCAAGTT	3089
Db	2821	TCTGTTGGGATGCAACGGAGCTGTCGTGGGCCATTCCTTTTGTGGGCGGCTCCCAAGTT	2880
Qy	3090	GTTATCTCTGGACGACCTTACGGCTTGGCTGGATCCTGCTTCCC CGCCGCGGTATTTGGGAG	3149
Db	2881	GTTATCTCTGGACGACCTTACGGCTTGGCTGGATCCTGCTTCCC CGCCGCGGTATTTGGGAG	2940
Qy	3150	CTGCTGCTCAAAATACCGAGAAAGTTCGACGCTGATTCCTCTCCACCAACACCTGGATGAG	3209
Db	2941	CTGCTGCTCAAAATACCGAGAAAGTTCGACGCTGATTCCTCTCCACCAACACCTGGATGAG	3000
Qy	3210	GCAGAGCTGCTGGAGACCGGTGCGCTGTGGTGGCAGGTGGCGCTGTGCTCTCTGTGGC	3269
Db	3001	GCAGAGCTGCTGGAGACCGGTGCGCTGTGGTGGCAGGTGGCGCTGTGCTCTCTGTGGC	3060
Qy	3270	TCCCCACTCTTCTCGCGCGCTCACCTGGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCC	3329
Db	3061	TCCCCACTCTTCTCGCGCGCTCACCTGGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCC	3120
Qy	3330	CGCTGCCCTTGACCAACAATGAGAAAGCTGACACTGACATGAGAGGCACTGTGGACAC	3389
Db	3121	CGCTGCCCTTGACCAACAATGAGAAAGCTGACACTGACATGAGAGGCACTGTGGACAC	3180
Qy	3390	AGGCAGAAAGAGAAATGCGACGCCAGGCGACAGAGTCCGCCACTCTCAGCTGCTGGCC	3449
Db	3181	AGGCAGAAAGAGAAATGCGACGCCAGGCGACAGAGTCCGCCACTCTCAGCTGCTGGCC	3240
Qy	3450	CTGGTACAGCACTGGGTGCCCGGGCACCGCTGCTGGAGGAGCTGCCACACGAGCTGGTG	3509
Db	3241	CTGGTACAGCACTGGGTGCCCGGGCACCGCTGCTGGAGGAGCTGCCACACGAGCTGGTG	3300
Qy	3510	CTGTTGTCGCCCTACACGGGTGCCCATGACGGCAGCTTTGGCCACACTCTTCCGAGAGCTA	3569
Db	3301	CTGTTGTCGCCCTACACGGGTGCCCATGACGGCAGCTTTGGCCACACTCTTCCGAGAGCTA	3360
Qy	3570	GACACGGGCTGGCGGAGCTGAGGCTCACCTGGCTACGGGATCTCCGACACCGCTCGAG	3629
Db	3361	GACACGGGCTGGCGGAGCTGAGGCTCACCTGGCTACGGGATCTCCGACACCGCTCGAG	3420
Qy	3630	GAGATCTTCTCTGAAGGTGGAGGAGTGTCTCGGACACAGATATGGAGGATGGCAGC	3689
Db	3421	GAGATCTTCTCTGAAGGTGGAGGAGTGTCTCGGACACAGATATGGAGGATGGCAGC	3480
Qy	3690	TGCGGGCAGCACTTATGACAGGCATTTCTGGCTAGACGTTAACCTTGGCTCAAGATG	3749
Db	3481	TGCGGGCAGCACTTATGACAGGCATTTCTGGCTAGACGTTAACCTTGGCTCAAGATG	3540
Qy	3750	CCGCCACGGAGACAGCGCTGGAGAACGGGGGACACGCTGGCTCAGCCCCCAGAGACTGAC	3809
Db	3541	CCGCCACGGAGACAGCGCTGGAGAACGGGGGACACGCTGGCTCAGCCCCCAGAGACTGAC	3600
Qy	3810	CAGGGCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTTGGGCACATGACCCCGCAGCAG	3869
Db	3601	CAGGGCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTTGGGCACATGACCCCGCAGCAG	3660
Qy	3870	CTCCAGGCCCTGCTTCTCAAGCGCTTTTCTGCTTGGCCCGCGACCGCGCGGCTGTTTC	3929

Db	3661		CTCCAGGCCCTGCTTCTCAAGACGCTTTTCTGCTTTGCCCGCGCAGCCGCCCGCGCTGTTC	3720
Qy	3930		GCCCAGATCGTGTGCTGCTGCCCTCTTTTGTGGCGCTGSCCCTCGTGTTACGCTCATCGTG	3989
Db	3721		GCCCAGATCGTGTGCTGCTGCCCTCTTTTGTGGCGCTGSCCCTCGTGTTACGCCTCATCGTG	3780
Qy	3990		CTCTCTTTGGGCACTAACCGGCTCTGCGGCTCAGTCCCAACCATGTACGGTGTCAAGTG	4049
Db	3781		CTCTCTTTGGGCACTAACCGGCTCTGCGGCTCAGTCCCAACCATGTACGGTGTCAAGTG	3840
Qy	4050		TCCTTCTTCAGTAGAGAGCGCCACAGGGACCTTGGAGTGGCCCGGCTGCTCGAGAGCGCTG	4109
Db	3841		TCCTTCTTCAGTAGAGAGCGCCACAGGGAGCCCTTGGAGTGGCCCGGCTGCTCGAGAGCGCTG	3900
Qy	4110		CTGCAGGAGCAGGACTTGGAGGAGCCCCAGTGCAGCATAGTCCACAGGTTCTCGGCA	4169
Db	3901		CTGCAGGAGGAGGACTTGGAGGAGCCCCAGTGCAGCATAGTCCACAGGTTCTCGGCA	3960
Qy	4170		CCAGAACTTCTTGTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTTGGACCCAGAGTCT	4229
Db	3961		CCAGAACTTCTTGTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTTGGACCCAGAGTCT	4020
Qy	4230		CCATCCCGACGCTGCCAGTGTAGCGAGCCCGGTGCCCCGGGCGCTGTCGCCGACTGCCCG	4289
Db	4021		CCATCCCGACGCTGCCAGTGTAGCGAGCCCGGTGCCCCGGGCGCTGTCGCCGACTGCCCG	4080
Qy	4290		GCTGCAGCTGCTGCTCCCTCGCCCGCCAGGAGTGACCGGCTCTGGGGAAGTGGTTTCA	4349
Db	4081		GCTGCAGCTGCTGCTCCCTCGCCCGCCAGGAGTGACCGGCTCTGGGGAAGTGGTTTCA	4140
Qy	4350		AACTGACGCGCGGAACCTTGTCTGACTTCTTGGTCAAGACTTACCGCGCGCTTGGTCGC	4409
Db	4141		AACTGACGCGCGGAACCTTGTCTGACTTCTTGGTCAAGACTTACCGCGCGCTTGGTCGC	4200
Qy	4410		CAGGGCTGAAGACTTAAGAAAGTGGGTGAATGAGGTCAAGTACGGAGGCTTCTCGCTGGG	4469
Db	4201		CAGGGCTGAAGACTTAAGAAAGTGGGTGAATGAGGTCAAGTACGGAGGCTTCTCGCTGGG	4260
Qy	4470		GGCGAGACCCAGGCTGCCCTCGGCGCAAGAGTGTGGCGCGCTCAGTGGAGGAGTTGTGG	4529
Db	4261		GGCGAGACCCAGGCTGCCCTCGGCGCAAGAGTGTGGCGCGCTCAGTGGAGGAGTTGTGG	4320
Qy	4530		GGCGTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTTGAAAACTTCACAGCC	4589
Db	4321		GGCGTGTGAGTCCCTGCTGCGGGGGCCCTCGACCGTGTCTTGAAAACTTCACAGCC	4380
Qy	4590		TGGGCTCACAGCTGTGATGTCTAGAGTCTCAAAGTCTGGTTTCAACCAAGAGGCTGG	4649
Db	4381		TGGGCTCACAGCTGTGACGCTCAGAGCAGTCTCAAGATCTGGTTTCAACCAAGAGGCTGG	4440
Qy	4650		CACCTCATGTGTGCTTTGTCAACCGAGCGAGCAACGCAATCTCTCGTGTCACTTCCCT	4709
Db	4441		CACCTCATGTGTGCTTTGTCAACCGAGCGAGCAACGCAATCTCTCGTGTCACTTCCCT	4500
Qy	4710		CCAGGCGCGGCCCGCCACGCGCCACAGCATCACACACTCAACCAACCCCTTGAACTCAC	4769
Db	4501		CCAGGCGCGGCCCGCCACGCGCCACAGCATCACACACTCAACCAACCCCTTGAACTCAC	4560
Qy	4770		AAGAGCAGCTGTGTAGGCTGTGACTGATGGCCCTCTCGGTGAGCGTCTCTGTTCTCAATC	4829
Db	4561		AAGAGCAGCTGTGTAGGCTGTGACTGATGGCCCTCTCGGTGAGCGTCTCTGTTCTCAATC	4620
Qy	4830		TGTGTGTCTTTGCCATGTCTTTGTCCGGCCAGCTTCACTCTGTCTCTCATTTGAGGAG	4889
Db	4621		TGTGTGTCTTTGCCATGTCTTTGTCCGGCCAGCTTCACTCTGTCTCTCATTTGAGGAG	4680
Qy	4890		CGAGTCAACCGAGCAGCACTTGCAGCTCATGGGGGGCGCTGTCCCGACCCCTTACTTGG	4949
Db	4681		CGAGTCAACCGAGCAGCACTTGCAGCTCATGGGGGGCGCTGTCCCGACCCCTTACTTGG	4740
Qy	4950		CTTGGCAACTTTTCTTGGGACATGTGTAACTTCTTGGTGGCCAGCATGTGTGTGTGCTC	5009

Db 4741 CTTGGCAACTTCTCTGGGACATGTGTAACTACTTGGTCCAGCATGCATCGTGTGCTC 4800
Qy 5010 ATCTTTCTGGCTTCCAGCAGAGGGCAATATGTGGCCCTTGCACCTGCTGCTCTCTG 5069
Db 4801 ATCTTTCTGGCTTCCAGCAGAGGGCAATATGTGGCCCTTGCACCTGCTGCTCTCTG 4860
Qy 5070 CTGTTGTCTACTGTATGTGTGTCATCACACCGCTCATGTACCCAGCCTCTCTTCTTC 5129
Db 4861 CTGTTGTCTACTGTATGTGTGTCATCACACCGCTCATGTACCCAGCCTCTCTTCTTC 4920
Qy 5130 TTCTCCGTGGCCAGCAGAGCCTATGTGTGTCACCTGCATATAACCTCTTTATTGGCATC 5189
Db 4921 TTCTCCGTGGCCAGCAGAGCCTATGTGTGTCACCTGCATATAACCTCTTTATTGGCATC 4980
Qy 5190 AATGGAAGCATGGCCACTTTGTGCTTGTAGCTCTTCTCTGATCAGAAAGCTGCGAGAGTG 5249
Db 4981 AATGGAAGCATGGCCACTTTGTGCTTGTAGCTCTTCTCTGATCAGAAAGCTGCGAGAGTG 5040
Qy 5250 AGCGGATCTTGAACAGGTCCTTCTTATCTTCCCCCACTTCTGCTTGGCCCGGGGCTC 5309
Db 5041 AGCGGATCTTGAACAGGTCCTTCTTATCTTCCCCCACTTCTGCTTGGCCCGGGGCTT 5100
Qy 5310 ATTGACATGTGGGAAACAGGCCATGGCTGATGCTTTTGAAGGCTTGGGAGACAGGCAG 5369
Db 5101 ATTGACATGTGGGAAACAGGCCATGGCTGATGCTTTTGAAGGCTTGGGAGACAGGCAG 5160
Qy 5370 TTCAGTCAACCTTGCCTGGAGGTGTGGCAAGAACCTCTTGGCCATGGTGAACAG 5429
Db 5161 TTCAGTCAACCTTGCCTGGAGGTGTGGCAAGAACCTCTTGGCCATGGTGAACAG 5220
Qy 5430 GGGCCCTCTTCTTCTTCACTACTGCTGACGACCGAAGCCAACTCTGCCACAG 5489
Db 5221 GGGCCCTCTTCTTCTTCACTACTGCTGACGACCGAAGCCAACTCTGCCACAG 5280
Qy 5490 CCCAGGTGAGGTCTCTGCCACTCTCGGAGAGAGACAGAGATGTAGCCCTGGAACGG 5549
Db 5281 CCCAGGTGAGGTCTCTGCCACTCTCGGAGAGAGACAGAGATGTAGCCCTGGAACGG 5340
Qy 5550 GAGCGGTGTTCNAGAGACCCAGGGGATGTGTGTGTGTGAGGAACTTGACCAAG 5609
Db 5341 GAGCGGTGTTCNAGAGACCCAGGGGATGTGTGTGTGTGAGGAACTTGACCAAG 5400
Qy 5610 GTATACCGTGGCAGAGATGCGAGCTGTGACCGCTTGTGCTGGGGATTTCCCTCTGT 5669
Db 5401 GTATACCGTGGCAGAGATGCGAGCTGTGACCGCTTGTGCTGGGGATTTCCCTCTGT 5460
Qy 5670 GAGTGTGTTGGCTGTCTGGGTGTGAATGAGCAGGGGAAGACGTCACGTTTTCGCATGGTG 5729
Db 5461 GAGTGTGTTGGCTGTCTGGGTGTGAATGAGCAGGGGAAGACGTCACGTTTTCGCATGGTG 5520
Qy 5730 ACGGGGACATTTGGCCAGAGGGGAGGCTGTGTGTGGAGGCCACAGCGTGGCCCGG 5789
Db 5521 ACGGGGACATTTGGCCAGAGGGGAGGCTGTGTGTGGAGGCCACAGCGTGGCCCGG 5580
Qy 5790 GAACCCAGTGTGGCACTCAGCATGGATCTGCCCTCAATCCGATGCCATCTTTGAG 5849
Db 5581 GAACCCAGTGTGGCACTCAGCATGGATCTGCCCTCAATCCGATGCCATCTTTGAG 5640
Qy 5850 CTGTGACGGGCCGACAGCACCTGAGCTGTGTCGCGCTGCGCGCTGTCGCGAGGCC 5909
Db 5641 CTGTGACGGGCCGACAGCACCTGAGCTGTGTCGCGCTGCGCGCTGTCGCGAGGCC 5700
Qy 5910 CAGTTGCCAGACCGCTGTGCTAGGCTGTGGCGTGTGGGATCTTCAATGGTACGCAGAC 5969
Db 5701 CAGTTGCCAGACCGCTGTGCTAGGCTGTGGCGTGTGGGATCTTCAATGGTACGCAGAC 5760
Qy 5970 CGGCTCAGGCACTACAGGGGGAACAAACGCAAGCTGGGAGCGCCCTGGCGCTG 6029
Db 5761 CGGCTCAGGCACTACAGGGGGAACAAACGCAAGCTGGGAGCGCCCTGGCGCTG 5820
Qy 6030 GTTGGGACCCAGCGGTGTGTTTCTGGAGAGCCGACCAAGGATGAGCCACCGCGC 6089
Db 5821 GTTGGGACCCAGCGGTGTGTTTCTGGAGAGCCGACCAAGGATGAGCCACCGCGC 5880

Qy 6090 CGCGCTTCTTTTGGAAACAGCCCTTTTGGCCGTGTGGAGGGCCGTTTCAGTGTGCTC 6149
Db 5881 CGCGCTTCTTTTGGAAACAGCCCTTTTGGCCGTGTGGAGGGCCGTTTCAGTGTGCTC 5940
Qy 6150 ACCTCCATAGCATGAGGAGTGTGAAGGCTCTGTCTCGCGCTAGCCATCATGTGTAAT 6209
Db 5941 ACCTCCATAGCATGAGGAGTGTGAAGGCTCTGTCTCGCGCTAGCCATCATGTGTAAT 6000
Qy 6210 GGCGGTTCGCTGTCTGGGACCCCGCAACATCTCAAGGGCAGATTTCGGCGGGTTCAC 6269
Db 6001 GGCGGTTCGCTGTCTGGGACCCCGCAACATCTCAAGGGCAGATTTCGGCGGGTTCAC 6060
Qy 6270 AACTGACCTTGGGTCGCGGTCGCGGCAAGGTCCAGCCGCGAGGCTTCTGTCGGCGCC 6329
Db 6061 AACTGACCTTGGGTCGCGGTCGCGGCAAGGTCCAGCCGCGAGGCTTCTGTCGGCGCC 6120
Qy 6330 GAGTTCCTTGGGTCGCGGTCGCGGAGGACATGAGAGGCGCTTCCAGCTTCCAGCTGCCG 6389
Db 6121 GAGTTCCTTGGGTCGCGGTCGCGGAGGACATGAGAGGCGCTTCCAGCTTCCAGCTGCCG 6180
Qy 6390 CCGGAGGGCGCTGCGCCCTTGGCGCGCTCTTTGGAGAGCTGCGCGTGCAACGCGCAGAG 6449
Db 6181 CCGGAGGGCGCTGCGCCCTTGGCGCGCTCTTTGGAGAGCTGCGCGTGCAACGCGCAGAG 6240
Qy 6450 CACGGGTGAGGACCTTTCCGTGAGGACGATGCTGGAGGAGTATCTTGTGTAATTC 6509
Db 6241 CACGGGTGAGGACCTTTCCGTGAGGACGATGCTGGAGGAGTATCTTGTGTAATTC 6300
Qy 6510 TCCAAAGCACCAGGGGAGGACGAGGACACCCAAAGAGCAGAGGAGGAGGAGTGGAGTG 6569
Db 6301 TCCAAAGCACCAGGGGAGGACGAGGACACCCAAAGAGCAGAGGAGGAGGAGTGGAGTG 6360
Qy 6570 GACCCCGCCAGGCTGACGACCCCAAAACGCGTCAGCGATTCCTCGATGACCTTAGC 6629
Db 6361 GACCCCGCCAGGCTGACGACCCCAAAACGCGTCAGCGATTCCTCGATGACCTTAGC 6420
Qy 6630 ACTGCGAGACTGTGCTGAGCCTCCCTCCCTGCGGGCCGCGGGAGGCCCTGGAA 6689
Db 6421 ACTGCGAGACTGTGCTGAGCCTCCCTCCCTGCGGGCCGCGGGAGGCCCTGGAA 6480
Qy 6690 TGGCAAGGGCAAGTAGAGTGTCTAGGAGCCCTGAGACTCAGGCTGGCAGAGGGCTGGTG 6749
Db 6481 TGGCAAGGGCAAGTAGAGTGTCTAGGAGCCCTGGACTCAGGCTGGCAGAGGGCTGGTG 6540
Qy 6750 CCTGAGAGAAATAAAGAGAGGCTGGAGAGAGCCGCTGCTGTGAAA 6797
Db 6541 CCTGAGAGAAATAAAGAGAGGCTGGAGAGAGCCGCTGCTGTGAAA 6588

RESULT 7

AX202218 LOCUS 6522 bp DNA linear PAT 30-AUG-2001
Sequence 1 from Patent WO0153490.

AX202218 DEFINITION

AX202218 ACCESSION

AX202218.1 GI:15392001

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Johns, M.A., Tafuri, S.R. and Wang, M.

Genes encoding abcl paralogs and the polypeptides derived therefrom

Patent: WO 0153490-A 1 26-JUL-2001;

WARNER-LAMBERT COMPANY (US)

Location/Qualifiers

1..6522

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Db 2101 TGGAGAGCCTGGCTCTGCTGGAGGAGCAGGGCGGAGGGCGGAGTGGCCACAACGTGGGC 2160
Qy 2370 ACCGGGCTACGGAGACGTCTTTCAGCCTGGCCGAGGTCTCTGGCCTTCTGCTGCTGGAC 2429
Db 2161 ACCGGGCTACGGAGACGTCTTTCAGCCTGGCCGAGGTCTCTGGCCTTCTGCTGCTGGAC 2220
Qy 2430 GCGGCGCTCTACGGGCTCGGCACCTGGTACTCTGGAAGCTGTGTGCCAGGCCAGTAGTCGGG 2489
Db 2221 GCGGCGCTCTACGGGCTCGGCACCTGGTACTCTGGAAGCTGTGTGCCAGGCCAGTAGTCGGG 2280
Qy 2490 ATCCCTGAACCATGGAAATTTTCCCTTTTCGAGGAGCTACTGGGTGCGGACCTCGGCCCCCC 2549
Db 2281 ATCCCTGAACCATGGAAATTTTCCCTTTTCGAGGAGCTACTGGGTGCGGACCTCGGCCCCCC 2340
Qy 2550 AAGAGTCCAGCCCTTGCCTCCACCCCGCTGGACCCCAAGAGTGTGTGTAAGAGGACCCG 2609
Db 2341 AAGAGTCCAGCCCTTGCCTCCACCCCGCTGGACCCCAAGAGTGTGTGTAAGAGGACCCG 2400
Qy 2610 CCGGCTGAGTCTCGGCGTATCCGTTTCGAGCTGTGGAGAGCGCTTTCTCTGAAGCCCG 2669
Db 2401 CCGGCTGAGTCTCGGCGTCTCGTTTCGAGCTGTGGAGAGCGCTTTCTCTGAAGCCCG 2460
Qy 2670 CAGCAGCCTCGGGGGCTCAGCTGGACTTTTACAGGGCCACATCACCGCTTCTCTG 2729
Db 2461 CAGCAGCCTCGGGGGCTCAGCTGGACTTTTACAGGGCCACATCACCGCTTCTCTG 2520
Qy 2730 GGGCACAACGGGGCGGCAAGACCAACACCTGTCTTACAGGGCCACATCACCGCTTCTCTG 2789
Db 2521 GGGCACAACGGGGCGGCAAGACCAACACCTGTCTTACAGGGCCACATCACCGCTTCTCTG 2580
Qy 2790 AGTGGTGGCTCTGCTTATCTCTGGGCGAGAGTCCGCTCCAGCATGGCGGCGCATCCGG 2849
Db 2581 AGTGGTGGCTCTGCTTATCTCTGGGCGAGAGTCCGCTCCAGCATGGCGGCGCATCCGG 2640
Qy 2850 CCCCACCTGGGCGCTCTGCTCTCAGTACAAAGCTGTGTTTGACATGTGACCGTGGACGAG 2909
Db 2641 CCCCACCTGGGCGCTCTGCTCTCAGTACAAAGCTGTGTTTGACATGTGACCGTGGACGAG 2700
Qy 2910 CAGTCTGGTCTTATGGGCGGCTGAGGGTCTGAGTCCGCTGTAGTGGGCCCGGAGCAG 2969
Db 2701 CAGTCTGGTCTTATGGGCGGCTGAGGGTCTGAGTCCGCTGTAGTGGGCCCGGAGCAG 2760
Qy 2970 GACCGTCTGTGAGGATGTGGGCTGCTTCCAAAGCAGAGTGTGAGACTTCGCCACCTC 3029
Db 2761 GACCGTCTGTGAGGATGTGGGCTGCTTCCAAAGCAGAGTGTGAGACTTCGCCACCTC 2820
Qy 3030 TCTGGTGGGATGCAACGGAAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 3089
Db 2821 TCTGGTGGGATGCAACGGAAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 2880
Qy 3090 GTTATCTGGACGAGCCTACGGCTGGCGTGGATCTCTGCTTCCCGCGCGGCTATTGGGAG 3149
Db 2881 GTTATCTGGACGAGCCTACGGCTGGCGTGGATCTCTGCTTCCCGCGCGGCTATTGGGAG 2940
Qy 3150 CTGCTGTCAAATAACGAGAGGTTCGACGCTGATCTCTCCACCAACACCTGGATGAG 3209
Db 2941 CTGCTGTCAAATAACGAGAGGTTCGACGCTGATCTCTCCACCAACACCTGGATGAG 3000
Qy 3210 GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTCTGTCTGGC 3269
Db 3001 GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTCTGTCTGGC 3060
Qy 3270 TCCCACTCTTCTGCGCGCTACCTGGGCTCCGGCTACTACTGAGCGTGGTGAAGGCC 3329
Db 3061 TCCCACTCTTCTGCGCGCTACCTGGGCTCCGGCTACTACTGAGCGTGGTGAAGGCC 3120
Qy 3330 CGCTCTGCCCTGACCAACATGAGAGCTGACACTGACATGAGGGGAGTGTGGACCC 3389
Db 3121 CGCTCTGCCCTGACCAACATGAGAGCTGACACTGACATGAGGGGAGTGTGGACCC 3180
Qy 3390 AGGCAGGAAAGAGAAATGGCAGCCAGGGCAGCAGAGTCCGCTCAGTCTGCTGGCC 3449
Db 3181 AGGCAGGAAAGAGAAATGGCAGCCAGGGCAGCAGAGTCCGCTCAGTCTGCTGGCC 3240

Qy 3450 CTGGTACAGCACTGGGTGCCCCGGGCAACGGCTGGTGGAGAGCTGCCACAGAGCTGGTG 3509
Db 3241 CTGGTACAGCACTGGGTGCCCCGGGCAACGGCTGGTGGAGAGCTGCCACAGAGCTGGTG 3300
Qy 3510 CTGGTCTGCTCCCTACACAGGCTGCCATGAGCGGAGCTTCCGCCACACTCTTTCGAGAGCTA 3569
Db 3301 CTGGTCTGCTCCCTACACAGGCTGCCATGAGCGGAGCTTCCGCCACACTCTTTCGAGAGCTA 3360
Qy 3570 GACAACGGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCAAGCTTCGAG 3629
Db 3361 GACAACGGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCAAGCTTCGAG 3420
Qy 3630 GAGATCTTCTGAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGAGAGGATGGCAGC 3689
Db 3421 GAGATCTTCTGAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGAGAGGATGGCAGC 3480
Qy 3690 TGGGGCAGCACCTATGACAGCACTTGTGGCTTAGACGTAAACCTTGGGGCTCAAGATG 3749
Db 3481 TGGGGCAGCACCTATGACAGCACTTGTGGCTTAGACGTAAACCTTGGGGCTCAAGATG 3540
Qy 3750 CCGCACAAGAGACAGGCTGGAGAAACGGGAAACAGCTTGGGTGAGCCCGCAGAGACTGAC 3809
Db 3541 CCGCACAAGAGACAGGCTGGAGAAACGGGAAACAGCTTGGGTGAGCCCGCAGAGACTGAC 3600
Qy 3810 CAGGGCTCTGGGCGACAGCGCTGGGCGGAGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3869
Db 3601 CAGGGCTCTGGGCGACAGCGCTGGGCGGAGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3660
Qy 3870 CTCAGGCGCTCTTCTCAAGCGCTTCTGCTTGGCGCGCAGCGCGCTGCTGCTGTTTC 3929
Db 3661 CTCAGGCGCTCTTCTCAAGCGCTTCTGCTTGGCGCGCAGCGCGCTGCTGCTGTTTC 3720
Qy 3930 GCGCAGATCTGCTGCTGCTGCTTGTGGGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3989
Db 3721 GCGCAGATCTGCTGCTGCTGCTTGTGGGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780
Qy 3990 CTTCTTTTCGGGCACTACCGGCTCTGCGCTCAGTCCACCATGATGAGTGTGCTGAGGTG 4049
Db 3781 CTTCTTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGATGAGTGTGCTGAGGTG 3840
Qy 4050 TCTTTCTTTCAGTGAGGACGCGCCAGGGGACCTTGGACGTGCCCGGCTGCTGAGGCGCTG 4109
Db 3841 TCTTTCTTTCAGTGAGGACGCGCCAGGGGACCTTGGACGTGCCCGGCTGCTGAGGCGCTG 3900
Qy 4110 CTGAGAGGACGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4169
Db 3901 CTGAGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960
Qy 4170 CCAGAGTTCCTGCTGAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCT 4229
Db 3961 CCAGAGTTCCTGCTGAGTGGCCAAAGGTCTTGGCCAGTGGCAACTGGACCCAGAGTCT 4020
Qy 4230 CCATCCCGAGCTGCCAGTGTAGCCAGCGGCTGCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 4289
Db 4021 CCATCCCGAGCTGCCAGTGTAGCCAGCGGCTGCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 4080
Qy 4290 GCTGAGCTGGTGGTCCCCCTCGGCCAGGAGGAGTACCGGCTCTGGGGAAGTGGTTCAG 4349
Db 4081 GCTGAGCTGGTGGTCCCCCTCGGCCAGGAGGAGTACCGGCTCTGGGGAAGTGGTTCAG 4140
Qy 4350 AACCTGACAGGCGGGAACCTGCTGACTTCTGCTCAAGACCTACCGCGGCTGGTGGCG 4409
Db 4141 AACCTGACAGGCGGGAACCTGCTGACTTCTGCTCAAGACCTACCGCGGCTGGTGGCG 4200
Qy 4410 CAGGGCTGGAAGACTAAGAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 4469
Db 4201 CAGGGCTGGAAGACTAAGAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 4260
Qy 4470 GGGCGAGACCCAGGCTGCTCGGCGCAAGAGTGGGCGGCTCAGTGGAGGAGTGGTGG 4529
Db 4261 GGGCGAGACCCAGGCTGCTCGGCGCAAGAGTGGGCGGCTCAGTGGAGGAGTGGTGG 4320

QY 4530 GGGTCTGAGTCCCTGCTGGCGGGGCCCTCGACCGGTGCTGTGAAACCTCAAGCC 4589
DB 4321 GGGTCTGAGTCCCTGCTGGCGGGGCCCTCGACCGGTGCTGTGAAACCTCAAGCC 4380
QY 4590 TGGGCTCACAGCCTGGATCTCAGGACAGTCTCAAGATCTGTTCAACAACAAAGGCTGG 4649
DB 4381 TGGGCTCACAGCCTGGATCTCAGGACAGTCTCAAGATCTGTTCAACAACAAAGGCTGG 4440
QY 4650 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACTGCCCC 4709
DB 4441 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACTGCCCC 4500
QY 4710 CCAGGCGGGGCCGCCACGCCACAGCATCAACAACCAACCCCTTGAACCTCAAC 4769
DB 4501 CCAGGCGGGGCCGCCACGCCACAGCATCAACAACCAACCCCTTGAACCTCAAC 4560
QY 4770 AAGGAGCAGCTCTGTAGGCTGCACTGATGGCTCTCGGTGGAAGTCTCGTCTCCATC 4829
DB 4561 AAGGAGCAGCTCTGTAGGCTGCACTGATGGCTCTCGGTGGAAGTCTCGTCTCCATC 4620
QY 4830 TGTGTGGTCTTTGCCATGTCCTTTGTCCGGCGAGCTTCACTCTTGTCTCTCATTTGAGGAG 4889
DB 4621 TGTGTGGTCTTTGCCATGTCCTTTGTCCGGCGAGCTTCACTCTTGTCTCTCATTTGAGGAG 4680
QY 4890 CGAGTCAACCGAGCCAAAGCACTGAGCTCATGGGGGCTGTCCCCACCCCTCTACTGG 4949
DB 4681 CGAGTCAACCGAGCCAAAGCACTGAGCTCATGGGGGCTGTCCCCACCCCTCTACTGG 4740
QY 4950 CTTGGCAACTTCTCTGGGACATGTGTAACTACTTTGGTGGCCAGCATGATCTGTGCTC 5009
DB 4741 CTTGGCAACTTCTCTGGGACATGTGTAACTACTTTGGTGGCCAGCATGATCTGTGCTC 4800
QY 5010 ATCTTTCTGGCCTTCCAGCAGAGGCAATGTGGCCCTGCGCAAACCTGCTCTCTCTG 5069
DB 4801 ATCTTTCTGGCCTTCCAGCAGAGGCAATGTGGCCCTGCGCAAACCTGCTCTCTCTG 4860
QY 5070 TGTGTTGTACTACTGTAATGGTGGTGCATCAACCGCTCATGTACCCAGCCTCTCTTCTTC 5129
DB 4861 TGTGTTGTACTACTGTAATGGTGGTGCATCAACCGCTCATGTACCCAGCCTCTCTTCTTC 4920
QY 5130 TTCTCGGTGCCAGCAGCCTATGTGTGCTCACTGATTAACCTCTTATTTGGCATC 5189
DB 4921 TTCTCGGTGCCAGCAGCCTATGTGTGCTCACTGATTAACCTCTTATTTGGCATC 4980
QY 5190 AATGAAGCATGGCCACCTTTGTGCTGAGCTCTTCTGTATCAAGAGCTGCAGGAGTG 5249
DB 4981 AATGAAGCATGGCCACCTTTGTGCTGAGCTCTTCTGTATCAAGAGCTGCAGGAGTG 5040
QY 5250 AGCGGATCTTTGAAACAGGCTCTCTTATCTTCCCACTTCTGCTTGGCGGGGGCTC 5309
DB 5041 AGCGGATCTTTGAAACAGGCTCTCTTATCTTCCCACTTCTGCTTGGCGGGGGCTC 5100
QY 5310 ATTGACATGGTGGGAAACAGGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGAG 5369
DB 5101 ATTGACATGGTGGGAAACAGGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGAG 5160
QY 5370 TTCCAGTCAACCTCCGCTGGAGGTGTGCGCAAGAACCTTCTGSCCATGCTGTATACAG 5429
DB 5161 TTCCAGTCAACCTCCGCTGGAGGTGTGCGCAAGAACCTTCTGSCCATGCTGTATACAG 5220
QY 5430 GGGCCCCCTCTTCTCTTCTTCACTACTGTGTGACAGCCGAAGCCAACTCTCTGCCACAG 5489
DB 5221 GGGCCCCCTCTTCTCTTCTTCACTACTGTGTGACAGCCGAAGCCAACTCTCTGCCACAG 5280
QY 5490 CCAGGCTGAGGTCTCTGCGCATCTCTGGAGAGGAGGACGAGGATGTAGCCGCTGAAACGG 5549
DB 5281 CCAGGCTGAGGTCTCTGCGCATCTCTGGAGAGGAGGACGAGGATGTAGCCGCTGAAACGG 5340
QY 5550 GAGCGGCTGCTCCAAAGGACCCACCCAGGGGATGTGTTGCTCTGAGGAACCTTGACCAAG 5609
DB 5341 GAGCGGCTGCTCCAAAGGACCCACCCAGGGGATGTGTTGCTCTGAGGAACCTTGACCAAG 5400
QY 5610 GTATACCGTGGGACAGGATGCCAGCTGTGTACCCGCTTGTGCTGGGATTTCCCCCTGGT 5669

DB 5401 GTATACCGTGGGACAGGATGCCAGCTGTGTGACCGCTTGTGCTGGGATTTCCCTGCTGT 5460
QY 5670 GAGTCTTTTGGCTCTGCTGGTGTGTAATGGAGCAGGAGGAGACCTGCACGTTTGCATGCTG 5729
DB 5461 GAGTCTTTTGGCTCTGCTGGTGTGTAATGGAGCAGGAGGAGACCTGCACGTTTGCATGCTG 5520
QY 5730 ACGGGGACACATTTGGCCAGCAGGCGCGAGGCTGTGCTGGCAGGCGACAGCCGTGGCCCCG 5789
DB 5521 ACGGGGACACATTTGGCCAGCAGGCGCGAGGCTGTGCTGGCAGGCGACAGCCGTGGCCCCG 5580
QY 5790 GAAACCAAGTGTGCGCACTCTAGCATGGGATATCTGCCCTCAATCCGATGCGATCTTTGAG 5849
DB 5581 GAAACCAAGTGTGCGCACTCTAGCATGGGATATCTGCCCTCAATCCGATGCGATCTTTGAG 5640
QY 5850 CTGCTGACGGGCGCGAGCACCTGAGCTGTGCTGGCGGCTGCGCGGTGTCCCGAGGCG 5909
DB 5641 CTGCTGACGGGCGCGAGCACCTGAGCTGTGCTGGCGGCTGCGCGGTGTCCCGAGGCG 5700
QY 5910 CAGGTTGCCACAGACCGCTGGCTCAGGCTGGCGCTCTGGGACTCTCATGTGTAACGAGAC 5969
DB 5701 CAGGTTGCCACAGACCGCTGGCTGGCGCTGGCGCTCTGGGACTCTCATGTGTAACGAGAC 5760
QY 5970 CGGCTGTGAGCACTTACAGCGGAGGAAACAAACGCAAGCTGGCGACCGCTCTGCGCTG 6029
DB 5761 CGGCTGTGAGCACTTACAGCGGAGGAAACAAACGCAAGCTGGCGACCGCTCTGCGCTG 5820
QY 6030 GTTGGGACCCAGCGCTGTTTCTGACGAGCCGACACAGGATGAGACCCAGCGCG 6089
DB 5821 GTTGGGACCCAGCGCTGTTTCTGACGAGCCGACACAGGATGAGACCCAGCGCG 5880
QY 6090 CGGCTCTCTTTGGAAACAGCCTTTTGGCGCTGGTGGCGGAGGCGCTTCACTGATGCTC 6149
DB 5881 CGGCTCTCTTTGGAAACAGCCTTTTGGCGCTGGTGGCGGAGGCGCTTCACTGATGCTC 5940
QY 6150 ACCTCCCATAGATGAGGAGTGTGAACGCTCTGCTCGCGCTTAGCCATCATGTGTAAT 6209
DB 5941 ACCTCCCATAGATGAGGAGTGTGAACGCTCTGCTCGCGCTTAGCCATCATGTGTAAT 6000
QY 6210 GGGCGTTTCCGCTGCTGGGAGCCGCGCAACATCTCAAGGCGAGATTCGGCGGGGCTAC 6269
DB 6001 GGGCGTTTCCGCTGCTGGGAGCCGCGCAACATCTCAAGGCGAGATTCGGCGGGGCTAC 6060
QY 6270 ACCTGACCTCTGGGTCGCCCGCCCAAGGCTCCAGCGCGAGCGGCTTCTGCTGGCGG 6329
DB 6061 ACCTGACCTCTGGGTCGCCCGCCCAAGGCTCCAGCGCGAGCGGCTTCTGCTGGCGG 6120
QY 6330 GAGTTCTCTGGTGGAGCTGCGGAGGCAATGAGAGCCGCTCGCTTCCAGCTGCGG 6389
DB 6121 GAGTTCTCTGGTGGAGCTGCGGAGGCAATGAGAGCCGCTCGCTTCCAGCTGCGG 6180
QY 6390 CCGGAGGCGCTGGCGCTGGCGGCTCTTTGGAGAGCTGGCGGCTGCAAGCGCGAGAG 6449
DB 6181 CCGGAGGCGCTGGCGCTGGCGGCTCTTTGGAGAGCTGGCGGCTGCAAGCGCGAGAG 6240
QY 6450 CACGCGTGGAGGACTTTTCCGTGAGCAGACGATGTGGAGGAGTATTTCTGTACTTC 6509
DB 6241 CACGCGTGGAGGACTTTTCCGTGAGCAGACGATGTGGAGGAGTATTTCTGTACTTC 6300
QY 6510 TCCAAAGGACCAAGGGGAGGACGAGGACACCCGAAAGACGAGGAGGAGGAGTGGAGTG 6569
DB 6301 TCCAAAGGACCAAGGGGAGGACGAGGACACCCGAAAGACGAGGAGGAGGAGTGGAGTG 6360
QY 6570 GACCCCGGCGAGGCTGCTGAGCAACCCGAAAGCGCTGAGGAGTCTCTGATGACCTTACG 6629
DB 6361 GACCCCGGCGAGGCTGCTGAGCAACCCGAAAGCGCTGAGGAGTCTCTGATGACCTTACG 6420
QY 6630 ACTGCGGAGACTGTGCTCTGAGCTTCCCTCCCTGCGGGGCGGCGGAGGAGGAGGAGG 6686
DB 6421 ACTGCGGAGACTGTGCTCTGAGCTTCCCTCCCTGCGGGGCGGCGGAGGAGGAGGAGG 6480
QY 6687 GAATGGCAAGGCAAGGTAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6727

Db	6481	GAATGGCAAGCGCAAGTAAATGCGTANGAGCCCTTGAAT	6521	
RESULT 8				
AX320364				
LOCUS		6432 bp	DNA	linear PAT 14-DEC-2001
DEFINITION		Sequence 3 from Patent WO0187978.		
ACCESSION		AX320364		
VERSION		AX320364.1	GI:17901744	
KEYWORDS				
SOURCE				
ORGANISM		Homo sapiens (human)		
REFERENCE				
AUTHORS		Glucksman, M. and Curtis, R.A.		
TITLE		38594, a novel human transporter and uses thereof		
JOURNAL		Patent: WO 0187978-A 3 22-NOV-2001;		
		Millennium Pharmaceuticals, Inc. (US)		
FEATURES				
source		Location/Qualifiers		
		1..6432		
		/organism="Homo sapiens"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:9606"		
CDS		1..->6432		
		/notes="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAD19392.1"		
		/db_xref="GI:17901745"		
		/translations="MAFWTQLMLLWKNFMYRRQPLLVELLWPLFLFTLVAVRHSH		
		PLLEHHECHFPNKLPSAGTVPMQLGLICNVNNTCPOLTPGEEPRGLNFNDLSVSR		
		LLADARTVLGGASAHRTLAGLGLIATLRAARSTAOPTKQSPFLPPMLDVAELLTS		
		LLRTESLGLAQAQPLSLLEAGDLAQELLALRLSLVELRALLORPGTSGPFLSLL		
		SEALCSVRGSSVTGSLNWEASDLMELVGPESALPDSSLPASCSELIGALDHP		
		LSRLWRRLKPLGLLPADPTFTRKLMQVNRFTFELTLRLDREVMELMGLPRIF		
		TFWDSNVAMQLRLQMODEGRQPRPGDRHMEALRSFLDPGSGYSWQDAHDVVG		
		HLVGTGRVTECLSLQLEAA PSEALVSRALOLLAEHRFWAGVFLGSDSDPTEH		
		PTDPLGPHVRIKIRMDIVVTRTKIRDRPWPDPAAADPLDLRLVYWGVSFYLQDLV		
		ERAARVLSGANPAGLILQOMPICYDDVDFURVLSRSLPFLFTLWVLSYVLTWKA		
		VREKTRLDTRWRAGLSLWLVLMFLGFLPILLSAALLVLKGLDILPYSHG		
		VVFLFAAFVATVTSFLLSAFFSRANLAAACGGLAYFSLYLPYLVLCVSWAMRDLPAQ		
		GRVAALLSPVAFGFCELSALLLEOGEAGQWNVGTRPTADVFSLAQVSGLLLDAA		
		LYGLATWYEAFCVPGQGIPEPNFPRRSYWCGRPPKPSAPCPTPLDPKVLVEEAP		
		PGLSPGVSRSLERKPSGPAPALRGLSLDFYQGHITAFLGHSAGKTTLSLSGLF		
		PPSGSGAFTLGHDRVMSMAIRPHLGVCPOINVLFDMLTVDEHWFIKRLKGLSAVW		
		QPQDRLLDVGLVSKVSQVTRHLSGMRKLSVAFAVGGSCQVWILDEPTAGVDPAS		
		RRGITWELLKYREGRTLILSTHLEDAELLGDRVAVAVGRLCCGSLPLRRLHSGG		
		YVLTLVKARLPLTTNKADTDMGSDVTRQEKNGSGQSVRGVTPQLLALVQHWVPGAR		
		LVSELPHLVLVLPYTGADHDSFATLPRDLTFLAELRLTGYSIDTSLSEIIFLVVE		
		ECADTDMDGSCGHLCTGIAGLDVTLRLKMPQETALENGEPAGSAPETDQSGPD		
		AVGRVQWALTRQQLQALLKRLFLARRSRRLGFIQIVLIPALFVGLALVSLIIVPFG		
		HYPALRLSPMYGAQVSFFSEDPAGDPGRARLLLEALLQEALEGPVQVSHSRFSAPF		
		VPAEAVKLVSNWTVTPESPSAPACQSRPGARLLPDCPAAAGPPPPQAVTSGSEVVQ		
		NLTGRNLSDFLVKTYTFLRVQGLTKWNVNRYVGFSGSLGRDPGLPSQGLGRSVEE		
		LWALLSPICGALDRVLKNTAVAHSLDAQDSLKIMWNKGMHSMVAFVNSASNAILR		
		AHLPPGARHAHSLITLNLPLNTKQLSEALMASVDVLSVICVFPFMSFPASFT		
		LVLIBERTRAKHLQMLGSLPTLYMGNFLWDMCNLYVPACIVLILFLAFQORAYVA		
		PANLPALLLLLLLYLQVSIPTLMPYPSGFSPVSP2AYVVLTCINLFTGINGSMAFLYE		
		LFSDQKLOEYSRLTKQVFLIFPHFCLGRGLIDMRNOAMADAPERLGDQFQSFPLWRE		
		VVGKLLAMVYIQLPFLITLQLHRSOLLPOPRVRSILPLLGEEDVAREVRVVGQ		
		ATQGVLDVLNIPKTVRGOMPAVDRLCLGIPGCEGGLLVGNVAGKTSFTFRVGTDT		
		LASGERVALGHSVAEPESAHLISMGYCPQSDAIPELLTGREHLELLRLRGVPEAQV		
		AQTAGSGLRGLVSWADRPAGTYSGGNKKRLATALLVGDPAVFLDEPTTQWPSA		
		RRLFLWNSLLAVRGWSVMLTSHSMECEALCSRLAIMVNGRFRCLGSPQHLKGRFAA		
		GHTLRLVPAARSOFAAFVAAEFPGSELREAHGRLRFLPPGRCALARVFGELAV		
		HGAHGVDFEVSQTMLEEVFLYFSKQDKDETEBQKEAGVGVDPAQGLQHPKRVSQ		
		FLDDPSPTAEIVL"		
ORIGIN				
Query Match	94.0%	Score 6398.4;	DB 6;	Length 6432;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 6421;	Conservative	0;	Mismatches	11;
			Indels	6;
			Gaps	1;

Db 1075 |||||CAGCCAGACCTGGAGGCGGGACCAATGGAGGCGCTCGCATCTCTTTCTGGACCCCTGGG 1134
Qy 1350 AGCGGTGGCTACAGCTGGGAGGACGACACGCTGATGTGGGACACCTGTGTGGGACGCTG 1409
Db 1135 AGCGGTGGCTACAGCTGGGAGGAGGACACGCTGATGTGGGACACCTGTGTGGGACGCTG 1194
Qy 1410 GCGCGAGTGA CGGAGTGCCTGCTCTTGGACAAGCTGGAGGCGGCACTCTCAGAGGAGCC 1469
Db 1195 GCGCGAGTGA CGGAGTGCCTGCTCTTGGACAAGCTGGAGGCGGCACTCTCAGAGGAGCC 1254
Qy 1470 CTGGTGTGCGGGGCGCTGCAACTGCTCGGGAACATCGATTTCTGGGCGGCGCTGCTCTTC 1529
Db 1255 CTGGTGTGCGGGGCGCTGCAACTGCTCGGGAACATCGATTTCTGGGCGGCGCTGCTCTTC 1314
Qy 1530 TTGGGACCTGAGGACTCTTTCAGACCCCAAGAGACCCCAAGAGCTGGGCGGCGGCG 1589
Db 1315 TTGGGACCTGAGGACTCTTTCAGACCCCAAGAGACCCCAAGAGCTGGGCGGCGGCG 1374
Qy 1590 CACGTGCGCATCAAAATCCGATGGACAATTGACGTGGTCAAGAGGACCAATAAGATCAGG 1649
Db 1375 CACGTGCGCATCAAAATCCGATGGACAATTGACGTGGTCAAGAGGACCAATAAGATCAGG 1434
Qy 1650 GACAGGTTTGGGACCTGGGCGGCGGAGGACCCGCTGACCGACCTGGGCGCTACGTGTGG 1709
Db 1435 GACAGGTTTGGGACCTGGGCGGCGGAGGACCCGCTGACCGACCTGGGCGCTACGTGTGG 1494
Qy 1710 GCGGCTCTGCTGTACTGCTCAAGACCTGTGGAGCGTGCAGCGGTCGCGGTGCTCAGCGGC 1769
Db 1495 GCGGCTCTGCTGTACTGCTCAAGACCTGTGGAGCGTGCAGCGGTCGCGGTGCTCAGCGGC 1554
Qy 1770 GCCAAACCCCGGCGGCGCTTACCTGCAGCAGATGCCCTATCCGCTGTATGTGACGAC 1829
Db 1555 GCCAAACCCCGGCGGCGCTTACCTGCAGCAGATGCCCTATCCGCTGTATGTGACGAC 1614
Qy 1830 GTGTTCTCTGCTGTGCTGAGCGGCTGCTGCGGCTCTTCTGACCGCTGGCTGGATCTAC 1889
Db 1615 GTGTTCTCTGCTGTGCTGAGCGGCTGCTGCGGCTCTTCTGACCGCTGGCTGGATCTAC 1674
Qy 1890 TCCGTGACACTGACAGTGAAGGCGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1949
Db 1675 TCCGTGACACTGACAGTGAAGGCGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734
Qy 1950 ATGCGCGCATGGGCTCAGCGCGGCTGCTGTGGCTAGGCTGGTGTCTTCTCAGCTGCCTC 2009
Db 1735 ATGCGCGCATGGGCTCAGCGCGGCTGCTGTGGCTAGGCTGGTGTCTTCTCAGCTGCCTC 1794
Qy 2010 GGGCGCTTCTGCTCAGCGCGGCTGCTGTGGTCTGAGCTGCTCAAGCTGGGAGACATCTC 2069
Db 1795 GGGCGCTTCTGCTCAGCGCGGCTGCTGTGGTCTGAGCTGCTCAAGCTGGGAGACATCTC 1854
Qy 2070 CCTACAGCACCCGGGGTGGTCTTCTGTTTGGAGGCTTCCGCGGTGGCCACGCTG 2129
Db 1855 CCTACAGCACCCGGGGTGGTCTTCTGTTTGGAGGCTTCCGCGGTGGCCACGCTG 1914
Qy 2130 ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCAACCTGGCTGGCGGCTGCGGC 2189
Db 1915 ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCAACCTGGCTGGCGGCTGCGGC 1974
Qy 2190 GGCCTGGCCTACTTCTCCCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2249
Db 1975 GGCCTGGCCTACTTCTCCCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
Qy 2250 CTGCGCGGGTGGCGGCTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2309
Db 2035 CTGCGCGGGTGGCGGCTGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
Qy 2310 TCGAGAGCTTGGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2369
Db 2095 TCGAGAGCTTGGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2154
Qy 2370 ACCGCGCTACGCGAGACGCTTTACGCTGCGCCAGAGTCTTGGGCTTCTGCTGCTGAC 2429

Db 2155 ACCCGGCTTACGCGACAGCGTCTTACGCTGGGCCAGGCTCTGTGCTCTGTGAC 2214
Qy 2430 GGGCGCTCTACGCGCTGCCACCTGTGTACCTGGAGCTGTGTGCCACGCGCAGTACGGG 2489
Db 2215 GGGCGCTCTACGCGCTGCCACCTGTGTACCTGGAGCTGTGTGCCACGCGCAGTACGGG 2274
Qy 2490 ATCTCTGAACCAATGGAAATTTTCTTTCGGAGGAGCTACTGTGTGGGAGCTCTGGGCCCCC 2549
Db 2275 ATCTCTGAACCAATGGAAATTTTCTTTCGGAGGAGCTACTGTGTGGGAGCTCTGGGCCCCC 2334
Qy 2550 AAGAGTTCAGCCCTTGGCCCCACCCGCTGGACCCAAAGGTCTCTGTGTAGAGAGGACCG 2609
Db 2335 AAGAGTTCAGCCCTTGGCCCCACCCGCTGGACCCAAAGGTCTCTGTGTAGAGAGGACCG 2394
Qy 2610 CCGGCTCTGAGTCTGGGCTATCCGCTTCGAGCCTGGAGAGCGCTTCTCTGGAGAGCGG 2669
Db 2395 CCGGCTCTGAGTCTGGGCTATCCGCTTCGAGCCTGGAGAGCGCTTCTCTGGAGAGCGG 2454
Qy 2670 CAGCAGCCCTGCGGGGCTCAGCCTGGACTTCTTACAGGGCCACATCACCGCCTTCTCTG 2729
Db 2455 CAGCAGCCCTGCGGGGCTCAGCCTGGACTTCTTACAGGGCCACATCACCGCCTTCTCTG 2514
Qy 2730 GGCACAAACGGGCGCGGCAAGACACACCTGTCCATCTTGAAGTGGCTCTTCCACACCC 2789
Db 2515 GGCACAGCGGGCGCGGCAAGACACACCTGTCCATCTTGAAGTGGCTCTTCCACACCC 2574
Qy 2790 AGTGTGGCTCTGCTTCTCAGTACAAACGCTGTGTGTGACATCTGACCGTGGACGAG 2909
Db 2635 CCCCACCTGGGCGCTGCTCTCAGTACAAACGCTGTGTGACATCTGACCGTGGACGAG 2694
Qy 2910 CAGCTGTGGTCTATGGGCGGCTGAAGGCTGAGTGGCGCTGTAGTGGGCGGCTTCCAGGAG 2969
Db 2695 CAGCTGTGGTCTATGGGCGGCTGAAGGCTGAGTGGCGCTGTAGTGGGCGGCTTCCAGGAG 2754
Qy 2970 GACGCTGTGTCAGGATGTGGGCTGCTCAGAGCAGAGTGTGACAGCTGCGCACCTC 3029
Db 2755 GACGCTGTGTCAGGATGTGGGCTGCTCAGAGCAGAGTGTGACAGCTGCGCACCTC 2814
Qy 3030 TCTGTGGGATGCAACCGAAGCTGTCTCGTGGCCATTTGCTTGTGGGCGGCTTCCAAAGTT 3089
Db 2815 TCTGTGGGATGCAACCGAAGCTGTCTCGTGGCCATTTGCTTGTGGGCGGCTTCCAAAGTT 2874
Qy 3090 GTTATCTGTGACAGCTTACCGGCTGAGTGTGCTTCCCGCGCGGCTTATTTGGGAG 3149
Db 2875 GTTATCTGTGACAGCTTACCGGCTGAGTGTGCTTCCCGCGCGGCTTATTTGGGAG 2934
Qy 3150 CTGCTGTCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCCACCACTGGATGAG 3209
Db 2935 CTGCTGTCTCAATACCGAGAGGTCGACGCTGATCTCTCCACCCACCACTGGATGAG 2994
Qy 3210 GCAGAGCTGTCTGGGAGACCGCTGTGTGGCTGAGTGGCGGCTTGTGTGTGTGTGTGTGTG 3269
Db 2995 GCAGAGCTGTCTGGGAGACCGCTGTGTGGCTGAGTGGCGGCTTGTGTGTGTGTGTGTGTG 3054
Qy 3270 TCCCCACTTCTTCTGCGGCTTCACTGGGCTTCCGGCTACTACTGACGCTGTGTGTGTGTGTG 3329
Db 3055 TCCCCACTTCTTCTGCGGCTTCACTGGGCTTCCGGCTACTACTGACGCTGTGTGTGTGTGTG 3114
Qy 3330 CGCTGCGGCTGACCCCAATGAGAGGCTGACACTGACATGGAGGGGAGTGTGTGTGTGTGTGTG 3389
Db 3115 CGCTGCGGCTGACCCCAATGAGAGGCTGACACTGACATGGAGGGGAGTGTGTGTGTGTGTGTG 3174
Qy 3390 AGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3449
Db 3175 AGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3234
Qy 3450 CTGTGTACAGCACTGGGTGCGCGGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3509
Db 3235 CTGTGTACAGCACTGGGTGCGCGGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3294

QY	3510	CTGGTGTGCCCTACACGGGTGCCCATGACGGCAGCTTTCGCCACA	CTCTCTCCGAGAGCTA	3569
DB				
	3295	CTGGTGTGCCCTACACGGGTGCCCATGACGGCAGCTTTCGCCACA	CTCTCCGAGAGCTA	3354
QY	3570	GACACGGCGCTGGCGGAGCTGAGGCTCACTGGGCTACGGGATCTCCGAC	CACGAGCTCGAG	3629
DB				
	3355	GACACGGCGCTGGCGGAGCTGAGGCTCACTGGGCTACGGGATCTCCGAC	CACGAGCTCGAG	3414
QY	3630	GAGATCTTCTGAAGTGGTGGAGAGTGTCTCGGACACAGATATGGAGGATGGCAGC		3689
DB				
	3415	GAGATCTTCTGAAGTGGTGGAGAGTGTCTCGGACACAGATATGGAGGATGGCAGC		3474
QY	3690	TGCGGCGACACCTATGTCACAGGCATTTGCTGGCCCTAGACGTAA	CCCTCGCGCTCAAGATG	3749
DB				
	3475	TGCGGCGACACCTATGTCACAGGCATTTGCTGGCCCTAGACGTAA	CCCTCGCGCTCAAGATG	3534
QY	3750	CCGCCACAGGAGACAGCGCTGGAAACGGGGAAACCACTGGGTGAGCC	CCGACAGACTGAC	3809
DB				
	3535	CCGCCACAGGAGACAGCGCTGGAAACGGGGAAACCACTGGGTGAGCC	CCGACAGACTGAC	3594
QY	3810	CAGGGCTCTGGGCACAGACGCGTGGGCGGGTACAGGGCTTGGGCACTG	ACCCGCCACGAG	3869
DB				
	3595	CAGGGCTCTGGGCACAGACGCGTGGGCGGGTACAGGGCTTGGGCACTG	ACCCGCCACGAG	3654
QY	3870	CTCCAGGCGCTGTCTTCAAGCGCTTCTTGCTTCCCGCGCGCACCCGCG	CGCGCTGTTC	3929
DB				
	3655	CTCCAGGCGCTGTCTTCAAGCGCTTCTTGCTTCCCGCGCGCACCCGCG	CGCGCTGTTC	3714
QY	3930	GCCAGATCTGTGCTTGTGCGCTCTTTTGTGGGCTTGCGCTCTCGTGT	TTCAGCGCTCATCGTG	3989
DB				
	3715	GCCAGATCTGTGCTTGTGCGCTCTTTTGTGGGCTTGCGCTCTCGTGT	TTCAGCGCTCATCGTG	3774
QY	3990	CCTCTTTTGGGCACTACCGGCTCTCGGGCTCAGTCCACCATGTACGGT	GTCTCAGGTTG	4049
DB				
	3775	CCTCTTTTGGGCACTACCGGCTCTCGGGCTCAGTCCACCATGTACGGT	GTCTCAGGTTG	3834
QY	4050	TCCCTTCTCAGTAGGACGCCCCAGGGGACCTTGAGCTGCCCCGGCTGT	CTCGAGGCGCTG	4109
DB				
	3835	TCCCTTCTCAGTAGGACGCCCCAGGGGACCTTGAGCTGCCCCGGCTGT	CTCGAGGCGCTG	3894
QY	4110	CTCAGGAGGCAGGACTGGAGGACCCCCAGTGCAGCATAGCTCCACAG	TGTTCTCGGCA	4169
DB				
	3895	CTCAGGAGGCAGGACTGGAGGACCCCCAGTGCAGCATAGCTCCACAG	TGTTCTCGGCA	3954
QY	4170	CCAGAGTTCTGTGTAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGAC	CCCCAGAGTCT	4229
DB				
	3955	CCAGAGTTCTGTGTAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGAC	CCCCAGAGTCT	4014
QY	4230	CCATCCCGAGCTGCCAGTGTAGCCAGCCGGTGC	CGCGCTGTCTCGCGCTGCGCG	4289
DB				
	4015	CCATCCCGAGCTGCCAGTGTAGCCAGCCGGTGC	CGCGCTGTCTCGCGCTGCGCG	4074
QY	4290	GCTCAGCTGTGTGCTCCCTCTCCGCCCCAGGACGTGACCGGCTCT	CGGGGAAGTGGTTTTCAG	4349
DB				
	4075	GCTCAGCTGTGTGCTCCCTCTCCGCCCCAGGACGTGACCGGCTCT	CGGGGAAGTGGTTTTCAG	4134
QY	4350	AACCTGACAGGCGGGAACCTGTCTGACTTCTGTGTCAAGACCTTAC	CCCGCGCTGTGTGCGC	4409
DB				
	4135	AACCTGACAGGCGGGAACCTGTCTGACTTCTGTGTCAAGACCTTAC	CCCGCGCTGTGTGCGC	4194
QY	4410	CAGGGCTTGAAGACTAAGAGTGGGTGAATGAGGTACAGTACGAGGCT	TCTCGCTGGGG	4469
DB				
	4195	CAGGGCTTGAAGACTAAGAGTGGGTGAATGAGGTACAGTACGAGGCT	TCTCGCTGGGG	4254
QY	4470	GGCCGAGACCCAGGCTGCCCTCGGGGCAAGAGTTGGGCGCTCAGT	TGGAGGAGTTGTGG	4529
DB				
	4255	GGCCGAGACCCAGGCTGCCCTCGGGGCAAGAGTTGGGCGCTCAGT	TGGAGGAGTTGTGG	4314
QY	4530	GCGTGTGTGATCCCTGTCTGTGGGGGGCCCTCGACCGTGTCTTGAAA	AAACCTCAGAGCC	4589
DB				
	4315	GCGTGTGTGATCCCTGTCTGTGGGGGGCCCTCGACCGTGTCTTGAAA	AAACCTCAGAGCC	4374

Qy		TGGGCTCACAGCCTGATGCTCAGACAGTCTCAAGATCTGTGTTCAAACAACAAAGGCGTG	4649
Dd		TGGGCTCACAGCCTGATGCTCAGGACAGTCTCAAGATCTGTGTTCAAACAACAAAGGCGTG	4434
Qy		CATCCATGTTGTGGCTTTGTCAAACCAGCACAGAAACGCATCTCTCGTGCCTCACTGCCCC	4709
Dd		CATCCATGTTGTGGCTTTGTCAAACCAGCACAGAAACGCATCTCTCGTGCCTCACTGCCCC	4494
Qy		CCAGGCCCGCCCGCCACAGCATCACCACACTCAACCAACCCCTTGAAACCTCACC	4769
Dd		CCAGGCCCGCCCGCCACAGCATCACCACACTCAACCAACCCCTTGAAACCTCACC	4554
Qy		AAGGAGAGCTGTCTGAGGCTGCACTGAATGGCTCTCTCGGTGAGCGTCTCTCTCATC	4829
Dd		AAGGAGAGCTGTCTGAGGCTGCACTGAATGGCTCTCTCGGTGAGCGTCTCTCTCATC	4614
Qy		TGTGTGGTCTTTTGGCATGTCTTTGTCCGGCCAGCTTCACTCTTGTCTCTCATTGAGGAG	4889
Dd		TGTGTGGTCTTTTGGCATGTCTTTGTCCGGCCAGCTTCACTCTTGTCTCTCATTGAGGAG	4674
Qy		CGAGTCAACCGACCAAGCACTGCAAGCTCATGGGGGGCTGTCCCACCACCTCTACTCTG	4949
Dd		CGAGTCAACCGACCAAGCACTGCAAGCTCATGGGGGGCTGTCCCACCACCTCTACTCTG	4734
Qy		CTTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCACGATGCATCGTGGTGCTC	5009
Dd		CTTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCACGATGCATCGTGGTGCTC	4794
Qy		ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTGCCCCAACCCTGCTCTCTTC	5069
Dd		ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTGCCCCAACCCTGCTCTCTTC	4854
Qy		CTGTGTCTACTACTGTATGGCTGGTGCATCACAACCGCTCATGTACCAGGCTCTTCTTC	5129
Dd		CTGTGTCTACTACTGTATGGCTGGTGCATCACAACCGCTCATGTACCAGGCTCTTCTTC	4914
Qy		TTCTCCGTGCCACACAGCCTATGTGGTCTCACCTGCATAAACCTCTTTATGGCATC	5189
Dd		TTCTCCGTGCCACACAGCCTATGTGGTCTCACCTGCATAAACCTCTTTATGGCATC	4974
Qy		AATGGAAGCATGGCCACCTTTGTGCTTAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG	5249
Dd		AATGGAAGCATGGCCACCTTTGTGCTTAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG	5034
Qy		AGCCGGATCTTGAAAACAGGCTCTTCTTATCTTCCCCACCTCTGTCTGGGCCCGGGGCTC	5309
Dd		AGCCGGATCTTGAAAACAGGCTCTTCTTATCTTCCCCACCTCTGTCTGGGCCCGGGGCTC	5094
Qy		ATTGACATGGTGGGAAACAGGCCATATGGCTGATGCTTTGAGCGCTGGGAGACAGGCAG	5369
Dd		ATTGACATGGTGGGAAACAGGCCATATGGCTGATGCTTTGAGCGCTGGGAGACAGGCAG	5154
Qy		TTCCAGTCAACCTTGCTGGGAGGTGTCGGCAAGAACTCTTGGCCATGGTGTATACAG	5429
Dd		TTCCAGTCAACCTTGCTGGGAGGTGTCGGCAAGAACTCTTGGCCATGGTGTATACAG	5214
Qy		GGGCCCTCTTCTCTTCTTCACTACTCTCTGACGACCGGAACCAACTCTCTGCCACAG	5489
Dd		GGGCCCTCTTCTCTTCTTCACTACTCTCTGACGACCGGAACCAACTCTCTGCCACAG	5274
Qy		CCAGGGTGAGGTCTCTGCCACTCTCTGGGAGAGGACGAGGATGTAGCCCGTGAACCG	5549
Dd		CCAGGGTGAGGTCTCTGCCACTCTCTGGGAGAGGACGAGGATGTAGCCCGTGAACCG	5334
Qy		GAGCGGTGTCCAAAGAGCCAACCGAGGGGATGTGTGTGTCTGAGGAACTTTGACCAAG	5609
Dd		GAGCGGTGTGTCCAAAGAGCCAACCGAGGGGATGTGTGTGTCTGAGGAACTTTGACCAAG	5394
Qy		GTATACCGTGGGCAGAGGATGCCAGCTGTGACCGCTTTGTGCTTGGGATTTCCCCCTGGT	5669
Dd		GTATACCGTGGGCAGAGGATGCCAGCTGTGACCGCTTTGTGCTTGGGATTTCCCCCTGGT	5454
Qy		GAGTGTTTTTGGGCTGTGGGTGTGAATGGACGAGGGAAGACGTCCACGTTTTCGACATGGT	5729

Db 5455 GAGTGTGTTGGGCTGCTGGGTGTGAATGGAGCAGGGAAGACGCTCCACGTTTTCGCATGTG 5514
Qy 5730 ACGGGGACACATTGGCCAGCAGGCGGAGGCTGTGCTGGCAGGCAACAGCGCTGCCCCGG 5789
Db 5515 ACGGGGACACATTGGCCAGCAGGCGGAGGCTGTGCTGGCAGGCAACAGCGTGGCCCCGG 5574
Qy 5790 GAACCCAGTGTGCGCACCTCAGCATGGGATACCTCCCTCAATCGATGCGCATCTTTGAG 5849
Db 5575 GAACCCAGTGTGCGCACCTCAGCATGGGATACCTCCCTCAATCGATGCGCATCTTTGAG 5634
Qy 5850 CTGTGACGGGCGGAGACCTGAGCTGCTTGGCGCGCTGCGCGGCTGCGCGGAGCC 5909
Db 5635 CTGTGACGGGCGGAGACCTGAGCTGCTTGGCGCGCTGCGCGGCTGCGCGGAGCC 5694
Qy 5910 CAGGTTGCCAGACCGCTGGCTCAGGCGCTGCGCGCTCTGGGACTCTCATGTACGACGAC 5969
Db 5695 CAGGTTGCCAGACCGCTGGCTGCGCGCTCTGGGACTCTCATGTACGACGAC 5754
Qy 5970 CGGCTGACGACCTCAGCGGAGGGAACAAACGCAAGCTGGCGACCGGCGCTGCGCTG 6029
Db 5755 CGGCTGACGACCTCAGCGGAGGGAACAAACGCAAGCTGGCGACCGGCGCTGCGCTG 5814
Qy 6030 GTTGGGACCCAGCGCTGGTCTTTCGACGAGCCGACCAAGGCGATGACCCAGCGCG 6089
Db 5815 GTTGGGACCCAGCGCGTGGTCTTTCGACGAGCCGACCAAGGCGATGACCCAGCGCG 5874
Qy 6090 CGGCGCTTCTTTGGAACAGCCTTTGCGCGTGGTGGGAGGCGCTTCAGTGATGCTC 6149
Db 5875 CGGCGCTTCTTTGGAACAGCCTTTGCGCGTGGTGGGAGGCGCTTCAGTGATGCTC 5934
Qy 6150 ACCTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGTAAT 6209
Db 5935 ACCTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGTAAT 5994
Qy 6210 CGGCGGTTCCGCTGCGGAGCGCGCAACATCTCAAGGCGAGATTCCGCGCGGCTCAC 6269
Db 5995 GGGCGGTTCCGCTGCGGAGCGCGCAACATCTCAAGGCGAGATTCCGCGCGGCTCAC 6054
Qy 6270 ACATGACCTTCCGCTGCGGAGCGCGCAAGGTTCCAGCGCGGCGCTTCTGTCGCGCC 6329
Db 6055 ACATGACCTTCCGCTGCGGAGCGCGCAAGGTTCCAGCGCGGCGCTTCTGTCGCGCC 6114
Qy 6330 GAGTTCCTGCGGTCGCGAGCTGCGGAGGACACATGAGGCGCGCTCGCTTCCAGCTGCGC 6389
Db 6115 GAGTTCCTGCGGTCGCGAGCTGCGGAGGACACATGAGGTCGCTCGCTTCCAGCTGCGC 6174
Qy 6390 CCGGAGGCGCTGCGCGCTGCGGCGCTCTTTGAGAGCTGCGGTCGACGCGCGAGAG 6449
Db 6175 CCGGAGGCGCTGCGCGCTGCGGCGCTCTTTGAGAGCTGCGGTCGACGCGCGAGAG 6234
Qy 6450 CACGCGTGGAGGACTTTTTCGCTGAGCCAGACGATGCTGGAGGAGGTATTTCTGTACTTC 6509
Db 6235 CACGCGTGGAGGACTTTTTCGCTGAGCCAGACGATGCTGGAGGAGGTATTTCTGTACTTC 6294
Qy 6510 TCCAGGACACAGGGAAGGACGAGGACACCGAAGCAGGAGGAGGAGGAGGAGGAGGAGG 6569
Db 6295 TCCAGGACACAGGGAAGGACGAGGACACCGAAGCAGGAGGAGGAGGAGGAGGAGGAGG 6354
Qy 6570 GACCCGCGCAGGCTGCGAGACCCCAACCGCTGACGAGTTCCTCGATGACCCCTAGC 6629
Db 6355 GACCCGCGCAGGCTGCGAGACCCCAACCGCTGACGAGTTCCTCGATGACCCCTAGC 6414
Qy 6630 ACTCGGAGACTGTGCTC 6647
Db 6415 ACTCGGAGACTGTGCTC 6432

RESULT 9
LOCUS AB055390 6027 bp mRNA linear PRI 11-AUG-2001
DEFINITION Homo sapiens ABCA7/ABCA-SSN mRNA for ABCA-SSN, complete cds.
ACCESSION AB055390

AB055390.1 GI:15042033
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Tanaka, A., Ikeda, Y., Abe-Dohmae, S., Arakawa, R., Sadanami, K.,
Kidera, A., Nakagawa, S., Nagase, T., Aoki, R., Kioka, N., Amachi, T.,
Yokoyama, S. and Ueda, K.
Human ABCA1 contains a large amino-terminal extracellular domain
homologous to an epitope of Sjogren's Syndrome
Biochem. Biophys. Res. Commun. 283 (5), 1019-1025 (2001)
21255283
11355874
2 (bases 1 to 6027)
Ueda, K. and Kioka, N.
Direct Submission
Submitted (02-FEB-2001) Kazumitsu Ueda, Kyoto University, Division
of Applied Life Sciences; Sakyu, Kyoto, Kyoto 606-8502, Japan
(E-mail:ueda@kais.kyoto-u.ac.jp, Tel:81-75-753-6105,
Fax:81-75-753-6104)
Location/Qualifiers
1. .6027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/cell_type="thymus"
1. .6027
/gene="ABCA7/ABCA-SSN"
1. .6027
/gene="ABCA7/ABCA-SSN"
/codon_start=1
/product="ABCA-SSN"
/protein_id="BAB62294.1"
/db_xref="GI:15042034"
/translation="MVCLGTGQSGAGPLVSVQHCPLVSPQESLGALGAQAEPLHS
LLEAEDLAQELALRLVELRALLQPRGTSGPLELLSEALCSVRGSPSTVGPLSNW
YEADLMELVQEPESALPDSSLSAPCELSGALDSDHPLSLWRLLPLILGLKFLFA
PDPTFKLMAQNRTEELTLDRVREWELMPRIETFMNDSNVAMLRLLQMD
EGRQPPGRGRDMEALRSLFDGSGYSQDAHADVGLVGTIGRVTESLSDKLEA
APSAALVRSALQLLAERHWAGVFLGDESDPTETPDLPGLGPHVRIKIMDIDV
QMPYCVDDVFLRVLRSILPLFLAWIYSVTLTKAVVREKTRLDRTWRMGLSR
AVLWGLFSLCLGPFLLSAILVLKGLDILPYSHGVPFLPAAFAVAVTQTSFL
SAFPRANLAAACGLAYFSLYLVLCVAMRDRPAGRVAAALLSPVATGFCESL
ALIEBQEGQAQHNHVGTRPTADVFLAQVSGLLDLDALYGLATWYLEAVCPQGYIP
EPNFPFRSYWCPGPPKSPAPCTPLDPKVLVEEAPPGLSPPGVSRLSKRFPSP
QPALRGSLDFYQGHITAFILHNGAGKTTLSILGLPFGSGGAPILGHVRSMAA
IRPHLGVCPQYVNLFDMLTVDEHVMFYGRLLGSLAAVPPRQDRLLQDVGLVSKQSV
TRHLSGQMRKLSVIAIFVGSQVYILDEPTAGVDPASRRGIWELLKXRGRTILS
THHLEAELGDRVAVVAGGRLCCGSPFLRRLHLSGGYILTKVRLPLITNEKADT
DMGSDVTRDKNGVSGSRVGTQLIALVQHWVPGARLVEELHELVLVLPYTGADH
GSEATLFRLELRLAEURLTGYGISTSLBIEIFLKVVECAADTDMEDSGCQHLCTG
IAGDVTLRKMPPEALENGEPAGSAPETDQSGDPADVGRVQGWALTROQLALL
KRFLARRRGLPAQIVLPALFVGLALVFSLLI VPRGHYPALRLSPMTMYCAOVSSFS
EDARGDPGARLLLEALLQEALEBPVQHSRSHSAPEVPAEVAKVLSGNWTSPRS
PACQSPGARLLPDCPAAGGPPPPQAVTGSBVQNLITGRNLSDFLVKTYFLVR
QKTRKWNVRYGKSLGDRGDLGSLGELGRSVBELWALLSPGLPGALDRVLKNL
TAWHSLDAQDSLKIFWNGRSHWAFVNRASNAI LRAHLPGRARHASTITINH
LNLTKQLEFAALMASSVDVLSICVVPFAMSFPASFTVLI ESRVTRAKHLQMGGL
SPTLYLGNFLMCMNYLVPACIVLILFIAQOQAYVAPANLPALELLALLLXGMSITP
LMTSPSFFSPSTAYVLTICINLIFIGINGSMATFVLELSDQKQVSRILKQVFLI
FPHCLGRGLIDMVRNADAPERLDGRQFSPRLRWEVGNKLLAMVIOQLPFLFT
LLIQHRSQLLPQPRVRLPGLLEGEDEVARERERVQGTQGVLTNRNLTKVLRGQR
MPAVDRCLGIPGCEGGLGVNGAGKTSFTFMVTDTLASRGEAVLAGHSVAREPSA
AHLMSGYCQSDAIFELLTGREHLELLARLVPEAOVATAGSGLARGLSWADRP
AGTSGGNEKELATALVGDPAVVFLEPTGMDPSARFLWNLAVVREGSVML
TSHSMECEALCSRLAIVNGRCFLCSQHLKGRPAAGHTLTLRVFAARQPAAFV
AAFPFGBELREAGHGRURFOLPPGRCALARVFGELAVHGAHGVDFSVTSQMLEEV
FLYFSQDQKQDETEQKEAGVGVDPAFGLQHPKRVSQFLDDPSTAEVTL"

ORIGIN

Db 2179 GCTTTCATCTCTGGGCCAAGAGTCCTGGCTCCAGCATGGCGCCATCCGCGCCCACTCGGGC 2238
Qy 2862 GTCTGTCTCAGTACAACTGTCTGTGTGACATGCTGACCGTGGAGCGAGCAGCTGTGGTTC 2921
Db 2239 GTCTGTCTCAGTACAACTGTCTGTGTGACATGCTGACCGTGGAGCGAGCAGCTGTGGTTC 2298
Qy 2922 TATGGGCGGCTGAAGGGTCTGAGTGCCTGTGAGTGGCCCCGAGCAGGACCGTCTGCTG 2981
Db 2299 TATGGGCGGCTGAAGGGTCTGAGTGCCTGTGAGTGGCCCCGAGCAGGACCGTCTGCTG 2358
Qy 2982 CAGGATGTGGGGCTGGTCTCCAAAGCAGAGTGTGACACTCGGCCACTCTCTGGTGGGATG 3041
Db 2359 CAGGATGTGGGGCTGGTCTCCAAAGCAGAGTGTGACACTCGGCCACTCTCTGGTGGGATG 2418
Qy 3042 CAA CGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTCGGAC 3101
Db 2419 CAA CGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTCGGAC 2478
Qy 3102 GAGCTACGGCTGGCGTGGATTCCTGCTTCCCGCCCGGTA TTTGGAGCTGCTGCTCAAA 3161
Db 2479 GAGCTACGGCTGGCGTGGATTCCTGCTTCCCGCCCGGTA TTTGGAGCTGCTGCTCAAA 2538
Qy 3162 TACCGAGAGCTCGACACCTGATCTCTCCACCCACACCTGGATGAGGCGAGAGCTGCTG 3221
Db 2539 TACCGAGAGCTCGACACCTGATCTCTCCACCCACACCTGGATGAGGCGAGAGCTGCTG 2598
Qy 3222 GGAGACCTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 3281
Db 2599 GGAGACCTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTGCCCACTCTTC 2658
Qy 3282 CTGGCCCTACCTGGGCTCGGCTACTACTGACGCTGTGGTGAAGGCCCGGCTGCCCTG 3341
Db 2659 CTGGCCCTACCTGGGCTCGGCTACTACTGACGCTGTGGTGAAGGCCCGGCTGCCCTG 2718
Qy 3342 ACCAACAATGAGAGGCTGACACTGACATGAGGGCAGTGTGGA CAC CAGGCGAGGAAAG 3401
Db 2719 ACCAACAATGAGAGGCTGACACTGACATGAGGGCAGTGTGGA CAC CAGGCGAGGAAAG 2778
Qy 3402 AAGAAATGACGACGAGGCGAGAGTGGCACTCTCAGCTGTGGCCCTGGTACAGCAC 3461
Db 2779 AAGAAATGACGACGAGGCGAGAGTGGCACTCTCAGCTGTGGCCCTGGTACAGCAC 2838
Qy 3462 TGGGTGCCCGGGGCA CGGCTGGTGGAGAGCTGCCACAGAGCTGGTGTGGTGTGCC 3521
Db 2839 TGGGTGCCCGGGGCA CGGCTGGTGGAGAGCTGCCACAGAGCTGGTGTGGTGTGCC 2898
Qy 3522 TACACGGGTGCCATGACGGCAGCTTTCGCCACA CTCTTCCGAGAGCTAGACACGCGGCTG 3581
Db 2899 TACACGGGTGCCATGACGGCAGCTTTCGCCACA CTCTTCCGAGAGCTAGACACGCGGCTG 2958
Qy 3582 GCGGAGCTGAGGCTCAGCTGGCTAGGGATCTCCGACACACCGCCCTCGAGGAGATCTTCCTG 3641
Db 2959 GCGGAGCTGAGGCTCAGCTGGCTAGGGATCTCCGACACACCGCCCTCGAGGAGATCTTCCTG 3018
Qy 3642 AAGGTGTGGAGGAGTGTGCTGCGGACACAGATATGAGGATGGCAGCTGGCGGCGAGCAC 3701
Db 3019 AAGGTGTGGAGGAGTGTGCTGCGGACACAGATATGAGGATGGCAGCTGGCGGCGAGCAC 3078
Qy 3702 CTATGCA CAGGCA TTGCTGGCTTGA CGTAA CCGTGGGCTCAAGATGCGGCCACAGGAG 3761
Db 3079 CTATGCA CAGGCA TTGCTGGCTTGA CGTAA CCGTGGGCTCAAGATGCGGCCACAGGAG 3138
Qy 3762 ACAGGCTGGAGAACGGGGAACAGCTGGGTGACGCCACAGAGACTGACAGGCTCTGGG 3821
Db 3139 ACAGGCTGGAGAACGGGGAACAGCTGGGTGACGCCACAGAGACTGACAGGCTCTGGG 3198
Qy 3822 CCAGACGCGCTGGGCGGCTTACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCGCTG 3881
Db 3199 CCAGACGCGCTGGGCGGCTTACAGGGCTGGGCACTGACCCGCGCAGCAGCTCCAGGCGCTG 3258
Qy 3882 CTTCTCAAGCGCTTTCTGCTTCCCGCGCGCAGCGCGCGCTGTTCGCCCAGATCGTG 3941

Db 3259 CTTCTCAAGCGCTTTCTGCTTCCCGCGCGCAGCGCGCGGCTGTTCGCCCCAGATCGTG 3318
Qy 3942 CTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTGCCTCTTTTCGGG 4001
Db 3319 CTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTGCCTCTTTTCGGG 3378
Qy 4002 CACTACCGGCTCTGCGGCTCAGTCCCAACCATGTA CGGTGTCTCAGGTGTCTTTCACT 4061
Db 3379 CACTACCGGCTCTGCGGCTCAGTCCCAACCATGTA CGGTGTCTCAGGTGTCTTTCACT 3438
Qy 4062 GAGGACCGCCAGGAGCCCTGGAGCTGCCCGGCTGTCCGAGGCGCTGTCGAGGAGCA 4121
Db 3439 GAGGACCGCCAGGAGCCCTGGAGCTGCCCGGCTGTCCGAGGCGCTGTCGAGGAGCA 3498
Qy 4122 GGA CTGAGGAGCGCCCGCAGTCAGCATAGCTCCCA CAGGTTCTCGGCA CCAAGTTCTCT 4181
Db 3499 GGA CTGAGGAGCGCCCGCAGTCAGCATAGCTCCCA CAGGTTCTCGGCA CCAAGTTCTCT 3558
Qy 4182 GCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCAA CTGGACCCAGAGTCTCCATGCCAGCC 4241
Db 3559 GCTGAAGTGGCCAAAGGTCTTGGCCAGTGGCAA CTGGACCCAGAGTCTCCATGCCAGCC 3618
Qy 4242 TGCCAGTGTAGCGACCGGCTGCCCGGCTGTGCTGCCGACTGCCCGCTGCAGCTGGT 4301
Db 3619 TGCCAGTGTAGCGACCGGCTGCCCGGCTGTGCTGCCGACTGCCCGCTGCAGCTGGT 3678
Qy 4302 GGTCCCTCCCGCCCGCAGGAGTGAACCGCTCTGGGGAAGTGTTCAGAACCTGCAGGC 4361
Db 3679 GGTCCCTCCCGCCCGCAGGAGTGAACCGCTCTGGGGAAGTGTTCAGAACCTGCAGGC 3738
Qy 4362 CGGAACCTGTCTGACTTTCTCGGTCAAGACCTACCGCGCCTGGTGGCGCAGGCTGAAG 4421
Db 3739 CGGAACCTGTCTGACTTTCTCGGTCAAGACCTACCGCGCCTGGTGGCGCAGGCTGAAG 3798
Qy 4422 ACTAAGAGTGGGTGAATGAGGTCAAGGTACGGAGCTTCTCGCTGGGGGGCGGAGACCA 4481
Db 3799 ACTAAGAGTGGGTGAATGAGGTCAAGGTACGGAGCTTCTCGCTGGGGGGCGGAGACCA 3858
Qy 4482 GGCTGCCCTCGGCGCAAGAGTTGGGCGGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGT 4541
Db 3859 GGCTGCCCTCGGCGCAAGAGTTGGGCGGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGT 3918
Qy 4542 CCCCTGCTGGCGGGGCCCTCGACCGTGTCTGA AAAA CCTCACAGCTGGGCTCACAGC 4601
Db 3919 CCCCTGCTGGCGGGGCCCTCGA CCGTGTCTGA AAAA CCTCACAGCTGGGCTCACAGC 3978
Qy 4602 CTGATGCTCAGGACAGCTCTCAAGATCTGGTTTCAACAA CAAAGGCTGCATCTCATGGTG 4661
Db 3979 CTGAGCGCTCAGGACAGTCTCAAGATCTGGTTTCAACAA CAAAGGCTGCATCTCATGGTG 4038
Qy 4662 GCCTTTGTCAACCGAGCGCAGCAACGCAATCTCGTGTCTCACTGCTGCCCGCCAGGCGCGGC 4721
Db 4039 GCCTTTGTCAACCGAGCGCAGCAACGCAATCTCGTGTCTCACTGCTGCCCGCCAGGCGCGGC 4098
Qy 4722 GGCACGCCCA CAGCATCA CCACTCAAC CCA CCGCTTGA AACCCTCA CCAAGSAGAGCTG 4781
Db 4099 GGCACGCCCA CAGCATCA CCACTCAAC CCA CCGCTTGA AACCCTCA CCAAGSAGAGCTG 4158
Qy 4782 TCTGAGGCTGCACTGATGGGCTCTCGTGTGAGAGCTCTCGTCTCCATCTGTGTGTCTTT 4841
Db 4159 TTTGAGGCTGCACTGATGGGCTCTCGTGTGAGAGCTCTCGTCTCCATCTGTGTGTCTTT 4218
Qy 4842 GCCATGCTCTTTGTC CGGCGCAGCTTCACTCTGTCTCTCA TTTGAGGAGCGAGTCA CCGA 4901
Db 4219 GCCATGCTCTTTGTC CGGCGCAGCTTCACTCTGTCTCTCA TTTGAGGAGCGAGTCA CCGA 4278
Qy 4902 GCCAAGCA CTTGCA GCTCAT TGGGGGCTGTCCCGCCAC CCGCTCTA CTGGCTTGGCACTTT 4961
Db 4279 GCCAAGCA CTTGCA GCTCAT TGGGGGCTGTCCCGCCAC CCGCTCTA CTGGCTTGGCACTTT 4338
Qy 4962 CTCTGGGACATGTGTAACTACTTGGTGCCAGCATGTGATCTGTGGTGTCTCATCTTTCTGGCC 5021
Db 4339 CTCTGGGACATGTGTAACTACTTGGTGCCAGCATGTGATCTGTGGTGTCTCATCTTTCTGGCC 4398

Qy	5022	TTCCAGCAGAGGGCATATGTGGCCCTTGCACAACTGCTGCTCTCTGCTGTGTGCTACTA	5081
Db	4399	TTCCAGCAGAGGGCATATGTGGCCCTTGCACAACTGCTGCTCTCTGCTGTGTGCTACTA	4458
Qy	5082	CTGTATGGCTGGTCGATCAACACGCTCATGTACCCAGCTCTCTTCTTCTCTCCGTGCCCC	5141
Db	4459	CTGTATGGCTGGTCGATCAACACGCTCATGTACCCAGCTCTCTTCTTCTCTCCGTGCCCC	4518
Qy	5142	AGCACGCCTATGTGGTGCCTCACTGCATAAACCTCTTTATTGGCATCAATGGAAGCATG	5201
Db	4519	AGCACGCCTATGTGGTGCCTCACTGCATAAACCTCTTTATTGGCATCAATGGAAGCATG	4578
Qy	5202	GCACCTTTGTGCTTTGAGCTCTTCTCTGTATCAGAACTGCAGGAGGTGAGCCGGATCTTG	5261
Db	4579	GCACCTTTGTGCTTTGAGCTCTTCTCTGTATCAGAACTGCAGGAGGTGAGCCGGATCTTG	4638
Qy	5262	AAACAGGTCTTCTTATCTTCCCCACTCTCTGCTTGGGCCGGGGCTCATTTGACATGGTG	5321
Db	4639	AAACAGGTCTTCTTATCTTCCCCACTCTCTGCTTGGGCCGGGGCTTATTGACATGGTG	4698
Qy	5322	CGNAACAGGCGCATGGCTGATGCCCTTTGAGCGCTTGGGAGACAGGACAGTTCCAGTACCC	5381
Db	4699	CGNAACAGGCGCATGGCTGATGCCCTTTGAGCGCTTGGGAGACAGGACAGTTCCAGTACCC	4758
Qy	5382	CTCGCTGGGAGGTGGTCGACAGAACTCTTGGCCATGTGTATACAGGGGCCCTCTCTTC	5441
Db	4759	CTCGCTGGGAGGTGGTCGACAGAACTCTTGGCCATGTGTATACAGGGGCCCTCTCTTC	4818
Qy	5442	CTTCTCTTCACTACTGCTGCACACCCGAAGCCAACTCTGCGCACAGCCAGGGTGAGG	5501
Db	4819	CTTCTCTTCACTACTGCTGCACACCCGAAGCCAACTCTGCGCACAGCCAGGGTGAGG	4878
Qy	5502	TCTCTGCACCTCTGGAGAGAGAGACGAGGATGTAGCCCGTGAAACCGGAGCGGGTGCTC	5561
Db	4879	TCTCTGCACCTCTGGAGAGAGAGACGAGGATGTAGCCCGTGAAACCGGAGCGGGTGCTC	4938
Qy	5562	CAGGAGCCACCCAGGGGGATGTTGGTGCTCAGGAACTTGACCAAGGTATACCGTGGG	5621
Db	4939	CAGGAGCCACCCAGGGGGATGTTGGTGCTCAGGAACTTGACCAAGGTATACCGTGGG	4998
Qy	5622	CAGAGGATGCCAGCTGTTTGACCGCTTGTGCTCGGGATTCGCCCTGCTGAGTGTTTGGG	5681
Db	4999	CAGAGGATGCCAGCTGTTTGACCGCTTGTGCTCGGGATTCGCCCTGCTGAGTGTTTGGG	5058
Qy	5682	CTGCTGGGTGTGAATGGAGCAGGGAAAGCTCCAGTTCGATGTGTGACGGGGGACACA	5741
Db	5059	CTGCTGGGTGTGAATGGAGCAGGGAAAGCTCCAGTTCGATGTGTGACGGGGGACACA	5118
Qy	5742	TTGCCACAGAGGGCGAGGCTGTCTGGCAGGCCACAGCGTGGCCCGGGAAACCCAGTGCT	5801
Db	5119	TTGCCACAGAGGGCGAGGCTGTCTGGCAGGCCACAGCGTGGCCCGGGAAACCCAGTGCT	5178
Qy	5802	CGCACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5861
Db	5179	CGCACCTCAGCATGGGATACCTGCCCTCAATCCGATGCCATCTTTGAGCTGCTGACGGGC	5238
Qy	5862	CGCAGCACTTGGAGCTGCTTTGGCGCCTTGGCGGTTGTCCGAGGCCCAAGTGTGCCAG	5921
Db	5239	CGCAGCACTTGGAGCTGCTTTGGCGCCTTGGCGGTTGTCCCGAGGCCCAAGTGTGCCAG	5298
Qy	5922	ACCGCTGGCTCAGCGCTTGGCGCTTGGGACTCTCATGTGTACGACACCGGCTGTGAGGC	5981
Db	5299	ACCGCTGGCTCAGCGCTTGGCGCTTGGGACTCTCATGTGTACGACACCGGCTGTGAGGC	5358
Qy	5982	ACCTACAGCGAGGGAAACAAACGCAAGCTGGCAGCGCCCTTGGCGCTTGGGGACCCA	6041
Db	5359	ACCTACAGCGAGGGAAACAAACGCAAGCTGGCAGCGCCCTTGGCGCTTGGGGACCCA	5418
Qy	6042	GCGGTGTGTTTCTGGACGACCGACACAGGCAATGACCCACAGCGCGCGCTTCTCTT	6101
Db	5419	GCGGTGTGTTTCTGGACGACCGACACAGGCAATGACCCACAGCGCGCGCTTCTCTT	5478

Qy	6102	TGGAACAGCCTTTTGGCCGCTGGTGC	CGGAGGGCCGTTTCACTGATGCTCA	CCTCCCATAGC	6161
Db	5479	TGGAACAGCCTTTTGGCCGCTGGTGC	CGGAGGGCCGTTTCACTGATGCTCA	CCTCCCATAGC	5538
Qy	6162	ATGGAGAGTGTGAAGCGCTCTGCT	CGCGCCTAGCCATCATGTGTGAATGG	CGCGTTCGCG	6221
Db	5539	ATGGAGAGTGTGAAGCGCTCTGCT	CGCGCCTAGCCATCATGTGTGAATGG	CGCGTTCGCG	5598
Qy	6222	TGCTCGGGCAGCCGCAACATCTCA	AGGGCAGATTTCGCGCGGGTACACA	CTGACCCCTG	6281
Db	5599	TGCTCGGGCAGCCGCAACATCTCA	AGGGCAGATTTCGCGCGGGTACACA	CTGACCCCTG	5658
Qy	6282	CGGCTGCCCGCCGCAAGTCTCC	AGCCGCGCAGCTTCTGTCGCGCG	CGAGTTCCTTGGG	6341
Db	5659	CGGCTGCCCGCCGCAAGTCTCC	AGCCGCGCAGCTTCTGTCGCGCG	CGAGTTCCTTGGG	5718
Qy	6342	TCGAGAGTTCGCGAGGACATATG	AGGCCGCTCTGGCTTCCAGCTGCG	CGGAGGGCGC	6401
Db	5719	TCGAGAGTTCGCGAGGACATATG	AGGCCGCTCTGGCTTCCAGCTGCG	CGGAGGGCGC	5778
Qy	6402	TGCGCCCTGCGCGCGCTTCTTGG	AGAGCTCGCGGTGCACGCGCAG	ACGCGTGGAG	6461
Db	5779	TGCGCCCTGCGCGCGCTTCTTGG	AGAGCTCGCGGTGCACGCGCAG	ACGCGTGGAG	5838
Qy	6462	GACTTTTCCGTGAGCCACAGATG	CTGTGGAGGAGTATTCTTGTACT	TCTTCCAAAGGACCAG	6521
Db	5839	GACTTTTCCGTGAGCCACAGATG	CTGTGGAGGAGTATTCTTGTACT	TCTTCCAAAGGACCAG	5898
Qy	6522	GGGAAGACAGAGACACCGAAG	ACGACGAAGGAGGAGGAGTGG	AGATGGAGACCCCGCCCA	6581
Db	5899	GGGAAGACAGAGACACCGAAG	ACGACGAAGGAGGAGGAGTGG	AGATGGAGACCCCGCCCA	5958
Qy	6582	GGCTCGACGACCCCAACCGCTC	AGCCAGTCTCTCGATCACCTAG	CACTGCGCGAGACT	6641
Db	5959	GGCTCGACGACCCCAACCGCTC	AGCCAGTCTCTCGATCACCTAG	CACTGCGCGAGACT	6018
Qy	6642	GTGCTCTGA	6650		
Db	6019	GTGCTCTGA	6027		
RESULT 10					
LOCUS		CQ719367		6324 bp DNA linear PAT 03-FEB-2004	
DEFINITION		Sequence 5301 from Patent WO02068579.			
ACCESSION		CQ719367			
VERSION		CQ719367.1 GI:42280224			
KEYWORDS					
SOURCE		Homo sapiens (human)			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS		1			
TITLE		Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W. Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof			
JOURNAL		Patent: WO 02068579-A 5301 06-SEP-2002;			
FEATURES		PE Corporation (NY) (US)			
source		Location/Qualifiers			
		1..6324			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			
ORIGIN					
Query Match		81.7%;		Score 5557.2; DB 6; Length 6324;	
Best Local Similarity		92.9%;		Pred. No. 0;	
Matches 6047; Conservative		0;		Mismatches 208; Indels 255; Gaps 11;	
Qy	210	ATGCGCTTCTGGACACAGCTGATG	CTCTGCTCTGGAAGAAATTTCA	TGTAATCGCCGAGA	269
Db	1	ATGCGCTTCTGGACACAGCTGATG	CTCTGCTCTGGAAGAAATTTCA	TGTAATCGCCGAGA	60

QY 270 CAGCGGTCAGCTCCTGGTGAATTTGCTGTGGCCCTCTCTTCTCTTCTTCTATCTGTG 329
DB |||||
61 CAGCGGTCAGCTCCTGGTGAATTTGCTGTGGCCCTCTCTTCTCTTCTTCTATCTGTG 120
QY 330 GCTGTTCCGCACTCCACCGCCCTGGAGCAACATGAATGCCATTCCTCCAAACAGCA 389
DB |||||
121 GCTGTTCCGCACTCCACCGCCCTGGAGCAACATGAATGCCATTCCTCCAAACAGCA 180
QY 390 CTGCCATCGGGGGCAGCGTGGCCCTGGCTCAGGGTCTCATCTGTAATGTGAACAAACC 449
DB |||||
181 CTGCCATCGGGGGCAGCGTGGCCCTGGCTCAGGGTCTCATCTGTAATGTGAACAAACC 240
QY 450 TGCTTTCCGCACTGACACCGGGCGAGAGCCCGGGCGCTGAGCAAATTCAAAGACTCC 509
DB |||||
241 TGCTTTCCGCACTGACACCGGGCGAGAGCCCGGGCGCTGAGCAAATTCAAAGACTCC 300
QY 510 CTGGTCTCCGGCTGTAGCCGATGCCGCACTGTGCTGGAGGGCCAGTGCACAGG 569
DB |||||
301 CTGGTCTCCGGCTGTAGCCGATGCCGCACTGTGCTGGAGGGCCAGTGCACAGG 360
QY 570 ACGTGGCTGGCTAGGGAAGCTGATCGCCACGCTGAGGGTGCACGAGCAACGCGCCAG 629
DB |||||
361 ACGTGGCTGGCTAGGGAAGCTGATCGCCACGCTGAGGGTGCACGAGCAACGCGCCAG 420
QY 630 CTTCAACCAA CCAAGCACTCTCACTGGAACCAACCCATGCTGGATGTGCGGAGCTGCTG 689
DB |||||
421 CTTCAACCAA CCAAGCACTCTCACTGGAACCAACCCATGCTGGATGTGCGGAGCTGCTG 480
QY 690 ACGTCACTGCTGGCAGCGAATCCCTGGGTTGGACCTGGGCAAGCCAGAGCCCTTG 749
DB |||||
481 ACGTCACTGCTGGCAGCGAATCCCTGGGTTGGACCTGGGCAAGCCAGAGCCCTTG 540
QY 750 CACAGCTTTGGAGCGCTGAGGACCTGGCCACAGGAGCTCCTGGCGCTGCGAGCCTG 809
DB |||||
541 CACAGCTTTGGAGCGCTGAGGACCTGGCCACAGGAGCTCCTGGCGCTGCGAGCCTG 600
QY 810 GTGAGCTTCGGGCACTGCTGAGAGACCCCGAGGAGCAGCGGCCCTCGAGTTGCTG 869
DB |||||
601 GTGAGCTTCGGGCACTGCTGAGAGACCCCGAGGAGCAGCGGCCCTCGAGTTGCTG 660
QY 870 TCAGAGCCCTTCGAGTGTAGGGGACCTTAGCAGCAGTGGGCCCCCTCCTCAACTGG 929
DB |||||
661 TCAGAGCCCTTCGAGTGTAGGGGACCTTAGCAGCAGTGGGCCCCCTCCTCAACTGG 720
QY 930 TACGAGGCTAGTGACCTGATGAGCTGTGGGAGGAGCAGAAATCCGCTGCGAGAC 989
DB |||||
721 TACGAGGCTAGTGACCTGATGAGCTGTGGGAGGAGCAGAAATCCGCTGCGAGAC 780
QY 990 AGCAGCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCAACCCGCTGTCC 1049
DB |||||
781 AGCAGCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCAACCCGCTGTCC 840
QY 1050 CGCTGCTCTGGAGACGCTGAAGCTCTGATCTCTGGGAAGCTACTCTTTGCAACAGAT 1109
DB |||||
841 CGCTGCTCTGGAGACGCTGAAGCTCTGATCTCTGGGAAGCTACTCTTTGCAACAGAT 900
QY 1110 ACACCTTTTACCGGAGCTCATGGCCAGGTGAGCCGAGCTTCGAGGAGCTCAACCTG 1169
DB |||||
901 ACACCTTTTACCGGAGCTCATGGCCAGGTGAGCCGAGCTTCGAGGAGCTCAACCTG 960
QY 1170 CTGAGGAGATGTCCGGAGGTGTGGAGATGTGGGACCCCGGATCTTCACTTATGAAC 1229
DB |||||
961 CTGAGGAGATGTCCGGAGGTGTGGAGATGTGGGACCCCGGATCTTCACTTATGAAC 1020
QY 1230 GACAGTTCCAAATGTGGCCATGCTGAGCGGCTCTCTGCAAGATGCAAGATGAAGGAAGG 1289
DB |||||
1021 GACAGTTCCAAATGTGGCCATGCTGAGCGGCTCTCTGCAAGATGCAAGATGAAGGAAGG 1047
QY 1290 CAGCCAGACCTGGAGGCGGGACCAACATGAGGCGCTTCCGATCCTTTCTGGACCTTGGG 1349
DB |||||
1048 ----- 1047

QY 1350 AGCGGTGCTACAGCTGGCAGGACGCA CACGCTGATGTGGGCA CTTGGTGGGCAAGCTG 1409
DB |||||
1048 ----- GTGTGGGGGTGCTGGGAGGTG 1071
QY 1410 GGCGAGTGACGGAGTGCTCTTGGACAAGCTGGAGGGGCA CTTCTCAGAGGACGAC 1469
DB |||||
1072 GGATGTGGCTCCCGTGCTCTTGGACAAGCTGGAGGGGCA CTTCTCAGAGGACGAC 1131
QY 1470 CTGTGTGCGGGGCCCTGCAACTGCTCGCGGAA CATCGATTTCTGGGCGCGGCTCGCTTC 1529
DB |||||
1132 CTGTGTGCGGGGCCCTGCAACTGCTCGCGGAA CATCGATTTCTGGGCGCGGCTCGCTTC 1191
QY 1530 TTGGGACCTGAGGACTCTTTCAGACCCACAGAGCA CCAACCCACAGACTTGGGCCCGGC 1589
DB |||||
1192 TTGGGACCTGAGGACTCTTTCAGACCCACAGAGCA CCAACCCACAGACTTGGGCCCGGC 1251
QY 1590 CACGTGGCATTCAAAATCCGCATGGACATTTGACCTGGTTCAGAGGACCAATAAGATCAGG 1649
DB |||||
1252 CACGTGGCATTCAAAATCCGCATGGACATTTGACCTGGTTCAGAGGACCAATAAGATCAGG 1311
QY 1650 GACAGGTTTGGGACCTTGGGCCAGCGGACCCCTTGAACCGACTTGGCTTACGTGTGG 1709
DB |||||
1312 GACAGGTTTGGGACCTTGGGCCAGCGGACCCCTTGAACCGACTTGGCTTACGTGTGG 1371
QY 1710 GGCGGCTTGTGTACTGTCAAGACTGTGTGAGCGGTG CAGCGGTCGCGTGTCTCAGCGGC 1769
DB |||||
1372 GGCGGCTTGTGTACTGTCAAGACTGTGTGAGCGGTG CAGCGGTCGCGTGTCTCAGCGGC 1431
QY 1770 GCCAACCCCGGGCGGGCTCTACTGTGAGCAGATGCCCTATCTCGTGTATGTGACGAC 1829
DB |||||
1432 GCCAACCCCGGGCGGGCTCTACTGTGAGCAGATGCCCTATCTCGTGTATGTGACGAC 1491
QY 1830 GTGTTCTCTGCTGTGTGAGCGGTGCTGCGGCTCTTCTTGAACGCTGGCTGATCTTAC 1889
DB |||||
1492 GTGTTCTCTGCTGTGTGAGCGGTGCTGCGGCTCTTCTTGAACGCTGGCTGATCTTAC 1551
QY 1890 TCCGTGACACTGACAGTGAAAGCGCTGTGCGGAGAA GGAAGCGCGCTGCGGACACC 1949
DB |||||
1552 TCCGTGACACTGACAGTGAAAGCGCTGTGCGGAGAA GGAAGCGCGCTGCGGACACC 1611
QY 1950 ATGCGCGCATGGGGCTCAGCGCGCGGTGCTCTGGCTAGCTTGGTTCTCAGCTTGCCTC 2009
DB |||||
1612 ATGCGCGCATGGGGCTCAGCGCGCGGTGCTCTGGCTAGCTTGGTTCTCAGCTTGCCTC 1671
QY 2010 GGCGCTTCTCTGCTCAGCGCGCTGCTGTGTTCTTGGTGCTCAAGCTGGGAGACATCTCTC 2069
DB |||||
1672 GGCGCTTCTCTGCTCAGCGCGCTGCTGTGTTCTTGGTGCTCAAGCTGGGAGACATCTCTC 1731
QY 2070 CCTTACAGCCACCCGGGCGTGTCTTCTGTTCTTGGCAGCTTTCGCGGTGGCCACGCTG 2129
DB |||||
1732 CCTTACAGCCACCCGGGCGTGTCTTCTGTTCTTGGCAGCTTTCGCGGTGGCCACGCTG 1791
QY 2130 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCAACCTTGGCTGGGCTGGCGGC 2189
DB |||||
1792 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGGCAACCTTGGCTGGGCTGGCGGC 1851
QY 2190 GGCTTGGCTTACTTCTCTCTCTCTCTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2249
DB |||||
1852 GGCTTGGCTTACTTCTCTCTCTCTCTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1911
QY 2250 CTGCGCGGGGTGGCGCGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2309
DB |||||
1912 CTGCGCGGGGTGGCGCGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1971
QY 2310 TGCAGAGCTTGGCTTCTGCTGGAAGGAGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2369
DB |||||
1972 TGCAGAGCTTGGCTTCTGCTGGAAGGAGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2031
QY 2370 ACCCGGCTCAGGAGAGCTTCTCAGGCTGGCCAGGCTCTCTGCGCTTCTGTGTGTGTGTGTGT 2429
DB |||||
2032 ACCCGGCTCAGGAGAGCTTCTCAGGCTGGCCAGGCTCTCTGCGCTTCTGTGTGTGTGTGTGT 2091
QY 2430 GCGGCGCTCTACGCGCTCGCCACCTGTGTACCTGTGAAGCTGTGTGTGTGTGTGTGTGTGTGT 2489

ACCESSION AX202220
VERSION AX202220.1 GI:15392002
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Johns, M.A., Tafuri, S.R. and Wang, M.
AUTHORS Genes encoding abci paralogs and the polypeptides derived therefrom
TITLE Patent: WO 0153490-A 3 26-JUL-2001;
JOURNAL WARNER-LAMBERT COMPANY (US)
FEATURES
Location/Qualifiers
source 1..5669
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 80.2%; Score 5458.2; DB 6; Length 5669;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5460; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 210 ATGSCCTTCTGGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCGGAGA 269
DB 1 ATGSCCTTCTGGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCGGAGA 60
QY 270 CAGCGGTCCAGCTCTCTGTCGAAATGCTGTGGCCTCTCTTCCTCTTTTCATCTGGTG 329
DB 61 CAGCGGTCCAGCTCTCTGTCGAAATGCTGTGGCCTCTCTTCCTCTTTTCATCTGGTG 120
QY 330 GCTGTTCCGCACTCCACCCGCGCTGGAGCACATGAATGCCACTTCCCAACAGCCCA 389
DB 121 GCTGTTCCGCACTCCACCCGCGCTGGAGCACATGAATGCCACTTCCCAACAGCCCA 180
QY 390 CTGCACTGCGCGGSCACCGTCTGCTGCTCAGGGTCTCATCTGTAATGTGAACAACACC 449
DB 181 CTGCACTGCGCGGSCACCGTCTGCTGCTCAGGGTCTCATCTGTAATGTGAACAACACC 240
QY 450 TGCTTTCCGCACTGACACCGGCGAGAGCCCGGCGCTGAGCAACTTCAACAGACTCC 509
DB 241 TGCTTTCCGCACTGACACCGGCGAGAGCCCGGCGCTGAGCAACTTCAACAGACTCC 300
QY 510 CTGCTCTCCGCGTCTAGCCGATGCCCGCACTGTCTGTGGAGGGGCGCACTGCCACAGG 569
DB 301 CTGCTCTCCGCGTCTAGCCGATGCCCGCACTGTCTGTGGAGGGGCGCACTGCCACAGG 360
QY 570 ACGTGTGCTGCGCTAGGCAAGCTGATGCCACGCTGAGGGCTGCACGACGACGCGGCCAG 629
DB 361 ACGTGTGCTGCGCTAGGCAAGCTGATGCCACGCTGAGGGCTGCACGACGACGCGGCCAG 420
QY 630 CCTCAACCAACCAAGCACTTCCACTGGAACCAACCCATGTGGATGTGCGGAGCTGCTG 689
DB 421 CCTCAACCAACCAAGCACTTCCACTGGAACCAACCCATGTGGATGTGCGGAGCTGCTG 480
QY 690 ACGTCACTGTGCGACGGAATTCCTGCGGTGTCACCTGGGCGACGAGGAGCCCTTG 749
DB 481 ACGTCACTGTGCGACGGAATTCCTGCGGTGTCACCTGGGCGACGAGGAGCCCTTG 540
QY 750 CACAGCTTTGTGGAGGCGCTGAGGACCTGCGCCAGGAGCTCCTGGGCGTGGCAGGCTG 809
DB 541 CACAGCTTTGTGGAGGCGCTGAGGACCTGCGCCAGGAGCTCCTGGGCGTGGCAGGCTG 600
QY 810 GTGAGCTTCGGGCACTGTGCGAGAGACCCCGAGGGACACGCGGCCCCCTTGGAGTTGCTG 869
DB 601 GTGAGCTTCGGGCACTGTGCGAGAGACCCCGAGGGACACGCGGCCCCCTTGGAGTTGCTG 660
QY 870 TCAGAGCCCTCTGCACTGTACGGGACCTAGCAGCAGTGTGGGCCCCCTTCAACTGG 929
DB 661 TCAGAGCCCTCTGCACTGTACGGGACCTAGCAGCAGTGTGGGCCCCCTTCAACTGG 720
QY 930 TAGAGGCTAGTACCTGATGGAGCTGGTGGGCGAGGACGAAATCCGCGCTGCCAGAC 989
DB 1801 GGGCCCTTCTGCTCAGCGCGCGCTGCTGGTTCGTGCTCAAGCTGGGGGACATCCTC 1860

DB 721 TACGAGGCTAGTGACCTGATGGAGCTGGTGGGCGAGGACCAAGATCCGCGCTGCCAGAC 780
QY 990 AGCAGGCTGAGCCCGCGCTGCTCGGAGCTGATTTGGAGCCCTGACAGCCACCCGCTGTCC 1049
DB 781 AGCAGGCTGAGCCCGCGCTGCTCGGAGCTGATTTGGAGCCCTGACAGCCACCCGCTGTCC 840
QY 1050 CGCCTGCTCTGGAGAGCCCTGGAAGCTCTGATCCTCGGGAAGCTACTCTTTTGACACAGAT 1109
DB 841 CGCCTGCTCTGGAGAGCCCTGGAAGCTCTGATCCTCGGGAAGCTACTCTTTTGACACAGAT 900
QY 1110 ACACCTTTTACCCCGGAAGCTCTATGCGCCAGGTGAACCGGACCTTTCGAGGAGCTCACCTG 1169
DB 901 ACACCTTTTACCCCGGAAGCTCTATGCGCCAGGTGAACCGGACCTTTCGAGGAGCTCACCTG 960
QY 1170 CTGAGGAGTGTCCGGAGGTGTGGGAGATGCTGGGAGCCCGGATCTTCACCTTCATGAAC 1229
DB 961 CTGAGGAGTGTCCGGAGGTGTGGGAGATGCTGGGAGCCCGGATCTTCACCTTCATGAAC 1020
QY 1230 GACAGTTTCCAATGCTGGCCATGCTGCAGCGCTCTCTGCAGATGCAGATGAAGGAAGG 1289
DB 1021 GACAGTTTCCAATGCTGGCCATGCTGCAGCGCTCTCTGCAGATGCAGATGAAGGAAGG 1080
QY 1290 CAGCCAGACCTCGAGGCGCGGACACATGGAGGCGCTCGCATCTCTTTTGGACCTTGGG 1349
DB 1081 CAGCCAGACCTCGAGGCGCGGACACATGGAGGCGCTCGCATCTCTTTTGGACCTTGGG 1140
QY 1350 AGCGGTGGCTACAGCTGGAGGACGACACAGCTGATGTGGGACCTTGTGGGACGCTG 1409
DB 1141 AGCGGTGGCTACAGCTGGAGGACGACACAGCTGATGTGGGACCTTGTGGGACGCTG 1200
QY 1410 GGCAGGAGTACAGGAGGCTCTCTTGGACAAGCTGGAGGCGCACCTTCAGAGGACGCC 1469
DB 1201 GGCAGGAGTACAGGAGGCTCTCTTGGACAAGCTGGAGGCGCACCTTCAGAGGACGCC 1360
QY 1470 CTGCTGTGCGGGCGCTGCAACTGCTCGCGGAAACATCGATTTCTGGGCGCGCTGCTTC 1529
DB 1261 CTGCTGTGCGGGCGCTGCAACTGCTCGCGGAAACATCGATTTCTGGGCGCGCTGCTTC 1320
QY 1530 TTGGGACCTGAGAGCTTTCAGACCCCAACAGAGACCCCAACCCAGACCTTGGGCGCGCG 1589
DB 1321 TTGGGACCTGAGAGCTTTCAGACCCCAACAGAGACCCCAACCCAGACCTTGGGCGCGCG 1380
QY 1590 CAGTGGCATCAAAATCCGATGGACATTTGAGTGTGTCACGAGGACCAATAAGATCAGG 1649
DB 1381 CAGTGGCATCAAAATCCGATGGACATTTGAGTGTGTCACGAGGACCAATAAGATCAGG 1440
QY 1650 GACAGGTTTGGGACCTTGGCCAGCGCGGACCCCTGTGACCGACCTTGCCTACGCTGG 1709
DB 1441 GACAGGTTTGGGACCTTGGCCAGCGCGGACCCCTGTGACCGACCTTGCCTACGCTGG 1500
QY 1710 GCGGCTTGTGTACCTGCAAGACCTGTGGAGCGTGCAGCCGCTCGCGTGTCTCAGCGCG 1769
DB 1501 GCGGCTTGTGTACCTGCAAGACCTGTGGAGCGTGCAGCCGCTCGCGCTGTCTCAGCGCG 1560
QY 1770 GCCAACCCCGCGCGCGCTTACCTGACAGATGCCCTATCCGTGTATGTGGAGAC 1829
DB 1561 GCCAACCCCGCGCGCGCTTACCTGACAGATGCCCTATCCGTGTATGTGGAGAC 1620
QY 1830 GTCTTCTGCTGTGTGAGCCGCTGCTGCGCTCTTCTGACGCTGGCTGATCTAC 1889
DB 1621 GTCTTCTGCTGTGTGAGCCGCTGCTGCGCTCTTCTGACGCTGGCTGATCTAC 1680
QY 1890 TCCGTGACACTGACAGTGAAGCCGTGTGGGAGAGAGAGAGCGGCTTGGGAGACCC 1949
DB 1681 TCCGTGACACTGACAGTGAAGCCGTGTGGGAGAGAGAGAGCGGCTTGGGAGACCC 1740
QY 1950 ATGCGCGCATGGGCTGACCGCGCGTGTCTGCTAGGCTGGTTCCTCAGCTGCTC 2009
DB 1741 ATGCGCGCATGGGCTGACCGCGCGTGTCTGCTAGGCTGGTTCCTCAGCTGCTC 1800
QY 2010 GGGCCCTTCTGCTCAGCGCGCGCTGCTGCTGCTGCTCAAGCTGGGGGACATCCTC 2069
DB 1801 GGGCCCTTCTGCTCAGCGCGCGCTGCTGCTGCTGCTCAAGCTGGGGGACATCCTC 1860

Db 4021 CCATCCCGACCTGCGCAGTGTAGCCAGCCGGTGCCTGCTGCCCGAGTCCCGC 4080
Qy 4290 GCTCAGCTGGTGGTCCCTCCGCCCCCAGGAGTACCGGCTCTGGGGAAGTGGTTCAG 4349
Db 4081 GCTCAGCTGGTGGTCCCTCCGCCCCCAGGAGTACCGGCTCTGGGGAAGTGGTTCAG 4140
Qy 4350 AACCTGACAGCGCGGAACCTGTCTGACTTCTGCTCAGACCTTACCGCGCTGCTGGCGC 4409
Db 4141 AACCTGACAGCGCGGAACCTGTCTGACTTCTGCTCAGACCTTACCGCGCTGCTGGCGC 4200
Qy 4410 CAGGCTGAAGACTAAGAGTGGTGAATAGAGTACGAGTACGAGGCTTCTCGCTGGGG 4469
Db 4201 CAGGCTGAAGACTAAGAGTGGTGAATAGAGTACGAGGCTTCTCGCTGGGG 4260
Qy 4470 GGCAGAGACCCAGGCTGCTCCCTCGGGCCAAAGTGTGGCCGCTCAGTGGAGAGTGTGG 4529
Db 4261 GGCAGAGACCCAGGCTGCTCCCTCGGGCCAAAGTGTGGCCGCTCAGTGGAGAGTGTGG 4320
Qy 4530 GCGTGTGATGTCCTGCTGCTGCGGGGCGCTCGACCGTGTCTGAAACCTCAGAGCC 4589
Db 4321 GCGTGTGATGTCCTGCTGCTGCGGGGCGCTCGACCGTGTCTGAAACCTCAGAGCC 4380
Qy 4590 TGGGCTCAGAGCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAACAAAGGCTGG 4649
Db 4381 TGGGCTCAGAGCTGGATGCTCAGGACAGTCTCAAGATCTGTTCAACAAACAAAGGCTGG 4440
Qy 4650 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCC 4709
Db 4441 CACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCC 4500
Qy 4710 CAGGCGCGCGCCGACCGCCACAGCATCACCACCTAACCAACCTGGAACCTCACC 4769
Db 4501 CAGGCGCGCGCGCCGACCGCCACAGCATCACCACCTAACCAACCTGGAACCTCACC 4560
Qy 4770 AAGGAGCAGCTGTCTGAGGCTGCACTGATGCGCTCTCGTGGAGCGTCTCGTCCATC 4829
Db 4561 AAGGAGCAGCTGTCTGAGGCTGCACTGATGCGCTCTCGTGGAGCGTCTCGTCCATC 4620
Qy 4830 TGTGTGCTTTTGGCATGTCTTTGTCCCGCCAGCTTCACTCTGTCTCATTTAGGAG 4889
Db 4621 TGTGTGCTTTTGGCATGTCTTTGTCCCGCCAGCTTCACTCTGTCTCATTTAGGAG 4680
Qy 4890 CGAGTCAACCGAGCAAGCACTGAGCTCATGCGGGGCGTGTCCCGCCAGCTTCACTG 4949
Db 4681 CGAGTCAACCGAGCAAGCACTGAGCTCATGCGGGGCGTGTCCCGCCAGCTTCACTG 4740
Qy 4950 CTGGGCAACTTCTCTGGGACATGTGTAACTACTTGTGTGCGGAGCATGCTGCTGCTC 5009
Db 4741 CTGGGCAACTTCTCTGGGACATGTGTAACTACTTGTGTGCGGAGCATGCTGCTGCTC 4800
Qy 5010 ATCTTTCTGGCCCTTCCAGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCCTG 5069
Db 4801 ATCTTTCTGGCCCTTCCAGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCCTG 4860
Qy 5070 CTGTGTCTACTGTATGCTGTGATCAGACCGCTCATGTATCCAGGCTCTCTTCTTC 5129
Db 4861 CTGTGTCTACTGTATGCTGTGATCAGACCGCTCATGTATCCAGGCTCTCTTCTTC 4920
Qy 5130 TTCTCCGTGCGCCAGCAGCAGCTATGTGTGCTCAGCTGCAATAAACCTTTATTGGCATC 5189
Db 4921 TTCTCCGTGCGCCAGCAGCAGCTATGTGTGCTCAGCTGCAATAAACCTTTATTGGCATC 4980
Qy 5190 AATGGAGCATGCGCCACTTTGTGCTTGTAGCTTCTTCTGATCAGAGCTGCGAGAGGTG 5249
Db 4981 AATGGAGCATGCGCCACTTTGTGCTTGTAGCTTCTTCTGATCAGAGCTGCGAGAGGTG 5040
Qy 5250 AGCCGATCTTGAACAGGCTTCTTATCTTCCCGCACTTCTGCTTGGCCGCGGGCTC 5309
Db 5041 AGCCGATCTTGAACAGGCTTCTTATCTTCCCGCACTTCTGCTTGGCCGCGGGCTC 5100
Qy 5310 ATTGACATGTGCGGAACAGGCGATGCGCTGATGCTTTTTCAGCGCTTGGGAGACAGGCG 5369

Db 5101 ATTGACATGTGCGGAACAGGCGCATGGCTGATGCTTTTGGAGCGCTTGGAGACAGGCG 5160
Qy 5370 TTCCAGTCACTCCCTGCTGCGGAGGTGGTGGCAAGAACTCTTGGCCATGGGTGATACAG 5429
Db 5161 TTCCAGTCACTCCCTGCTGCGGAGGTGGTGGCAAGAACTCTTGGCCATGGGTGATACAG 5220
Qy 5430 GGGCCCTCTTCTCTCTCTTCTTCACTACTGCTGACAGCACCGAAGCCAACTCTCTGCCACAG 5489
Db 5221 GGGCCCTCTTCTCTCTCTTCTTCACTACTGCTGACAGCACCGAAGCCAACTCTCTGCCACAG 5280
Qy 5490 CCAGGCTGAGGTCTCTGCGCACTCTCTGGGAGAGGAGGAGGATGTAGCCCTGAAACGG 5549
Db 5281 CCAGGCTGAGGTCTCTGCGCACTCTCTGGGAGAGGAGGAGGATGTAGCCCTGAAACGG 5340
Qy 5550 GAGCGGCTGCTCAAGAGAGCCACCAAGGGGATGTGTGCTGAGGAACTTGAACAG 5609
Db 5341 GAGCGGCTGCTCAAGAGAGCCACCAAGGGGATGTGTGCTGAGGAACTTGAACAG 5400
Qy 5610 GTATACCGTGGGAGAGGATGCGAGCTGTGACCGCTTGTGCTGGGATTTCCCTCGGT 5669
Db 5401 GTATACCGTGGGAGAGGATGCGAGCTGTGACCGCTTGTGCTGGGATTTCCCTCGGT 5460
Qy 5670 GAG 5672
Db 5461 GAG 5463

RESULT 12

AX644625
LOCUS AX644625
DEFINITION Sequence 7 from Patent WO02099108.
ACCESSION AX644625
VERSION AX644625.1 GI:28610639
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Shutter,J. and Ulias,L.
Atp-binding cassette transporter-like molecules and uses thereof
Patent: WO 02099108-A 7 12-DEC-2002;
Amgen, Inc. (US)

FEATURES

Location/Qualifiers
1..4653
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..4653
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD68133.1"
/db_xref="GI:28610640"
/translation="MVCLGTGOSAGPLVSVQNHCPGLSPQESLGLALGQAQBPPLHS
LLEAGDLAEILLALRLSELVELRALLORPRGTGPLELSEALCSVRGPSSTVGPLNHW
YEASDLNVLVQEPESALPDSSLPACSELIHALDHPRLRLRLRLKPLIKLLPLA
POTPTFKLMAQVNRTEBELTLRDVREVMELGPKRIFTMNDSSNVAMQLRLQMD
EGRQPRPGRDHMEALRFLDPGSGVSNQDAHDVHLVGLTGRVTRIKIRMDIDV
ASEARALSRALAEHRFWAGVFLGPEVSDPDTEHTPTDLPGLSGHVRIRIKIRMDIDV
VTRTKIRDRFWDPGPADPLTLRLVYVWGFFVYLQDLVERAAVRLSGANPRAGLYQ
OMPYCVDVFLVRLSLPLFLTLAWIVSVTLTKAVVREKETRLRDTMRAMGLSR
AVLWGLFSLCGLPFLLSALLVLKGLDILPYSHPGVFLFLAFAVATVTSLL
SAFFSRANLAAGCHGLAYFSLYLPLVCVAMRDLRPAAGRAASLSPVAFGCGSL
ALLEEGEQAQHWNVTRPTADVLSAQVSLGLLDLDAALYGLATWLEAVCPQYGLP
EPWNPFRSYWCPKSPAPCTPLDPKVLVEEAPPLSPGVSVRSLEKRFPGSP
QPALRGLSDYQGHITAFLHNGAGTKTTLSILSGLFPSSGSAFILDHVRSSMAA
IPLHGLVCPQYNVLFMLTVDEHVPYGLKGLSAAVVGPEDRLQDLVLSKQSVQ
THLSCGMOKLSVAIAFVGGSQVILDEPTAGVPASRGIIWELLIKVREGTLLLS
THLHDAELLGDRVAVVAGRLCCPLFLRRHLGSGYFLTLVKARLPLVTNEKADT
DMGVSVDTRFKKNGSQSRVTPQLLALVQHWVPGARLVEELPHELVLVLPYTGADH
GSPATLREKDLAEHLRTGYISDTLEEIPFKVVECAADTMEDTSCGQHLCTG
IAGLDVTLKMPPOETALENGEPAGSAPETDQSGDPAVGRVQGWALTTRQQLALL
KRFLLARRSRGLFAQIVLPALFVGLALVFSLLVPPFGHYPALRLSPTMYQVSPFS

CDS

EDAPDGRARLELLEALLOEAGLEBEPVOHSHRPSAPEVPAEVAKVILASGNWBTBESPS
PACQSCPGARLLLPDCFAAAGGPPPOAVTSGSEVQVQLTGRNLSDFLVKTYPLRLVR
GGLKTKWNEVRYGFSIGRDPGLPSQBELWALLSPFGCAGGACAAATAAGATCAGGGACAGGTTTGG 1661
979 AAAATCCGCATGGA CATTTGACGTGGTACAGAGGACAAATAAGATCAGGGACAGGTTTGG 1038
1662 GACCTGCGCCAGCGCGGACCCCCCTGACCGACCTGCGCTAGCTGTGGGGCGGCTTCGTG 1721
1039 GACCTGCGCCAGCGCGGACCCCCCTGACCGACCTGCGCTAGCTGTGGGGCGGCTTCGTG 1098
1722 TACCTGCAAGACCTGGTGGAGCGTGCAGCCGCTCCGCGTGTCTCAGCGGCGCCAAACCCCGG 1781
1099 TACCTGCAAGACCTGGTGGAGCGTGCAGCCGCTCCGCGTGTCTCAGCGGCGCCAAACCCCGG 1158
1782 GCGGCGCTCTACCTGACGAGATGCCCTATCCGCTGTATGTGGACGACGTGTTCCTGCGT 1841
1159 GCGGCGCTCTACCTGACGAGATGCCCTATCCGCTGTATGTGGACGACGTGTTCCTGCGT 1218
1842 GTGCTGAGCCCGTGGTGGAGCGTGCAGCCGCTCTTCTGACGCTGGGCTTGGATCTACTCCGTGACACTG 1901
1219 GTGCTGAGCCCGTGGTGGAGCGTGCAGCCGCTCTTCTGACGCTGGGCTTGGATCTACTCCGTGACACTG 1278
1902 ACAGTGAAGGCGGTGGTGGAGGAGGAGACGCGGCTGCGGGACACCATGCGCGCCATG 1961
1279 ACAGTGAAGGCGGTGGTGGAGGAGGAGACGCGGCTGCGGGACACCATGCGCGCCATG 1338
1962 GGGCTCAAGCGCGGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2021
1339 GGGCTCAAGCGCGGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1398
2022 CTGAGCGCGGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2081
1399 CTGAGCGCGGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1458
2082 CCGGCGGTGGTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2141
1459 CCGGCGGTGGTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1518
2142 CTGCTCAGCGCTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2201
1519 CTGCTCAGCGCTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1578
2202 TTCTCCCTCTACCTGCGCTGAGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2261
1579 TTCTCCCTCTACCTGCGCTGAGCTGCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1638
2262 GCGGCGGTGGTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 2321
1639 GCGGCGGTGGTTCCTGCTGAGCTAGGCTGGTTCCTCAGCTGCTCGGGCCCTTCCTG 1698
2322 GCTCTGCTGAGGAGCGGCGAGGCGCGAGTGGCAACAAGTGGGCGCCGCGCTTACG 2381
1699 GCTCTGCTGAGGAGCGGCGAGGCGCGAGTGGCAACAAGTGGGCGCCGCGCTTACG 1758
2382 GCAGAGCTTTCAGCTGCGGCGGCGAGTGGTTCCTGCTGAGCTAGGCGCGCTTCTAC 2441
1759 GCAGAGCTTTCAGCTGCGGCGGCGAGTGGTTCCTGCTGAGCTAGGCGCGCTTCTAC 1818
2442 GCGCTCGCCACTGCTGAGAGCTGCTGAGGCGCGAGTGGCAACAAGTGGGCGCCGCGCTTACG 2501
1819 GCGCTCGCCACTGCTGAGAGCTGCTGAGGCGCGAGTGGCAACAAGTGGGCGCCGCGCTTACG 1878
2502 TGGAAATTTCTTTTCGAGGAGCTTCTGCTGAGGCGCGAGTGGGCGCCGCGCTTACG 2561
1879 TGGAAATTTCTTTTCGAGGAGCTTCTGCTGAGGCGCGAGTGGGCGCCGCGCTTACG 1938
2562 CTTGCGCCACTGCTGAGAGCTGCTGAGGCGCGAGTGGGCGCCGCGCTTCTGAGT 2621
1939 CTTGCGCCACTGCTGAGAGCTGCTGAGGCGCGAGTGGGCGCCGCGCTTCTGAGT 1998
2622 CTTGCGCTTCTGAGAGCTGCTGAGGCGCGAGTGGGCGCCGCGCTTCTGAGT 2681
1999 CTTGCGCTTCTGAGAGCTGCTGAGGCGCGAGTGGGCGCCGCGCTTCTGAGT 2058
2682 CCGGCGGTCTGAGCTTCTTACAGGCGCGAGTGGGCGCCGCGCTTCTGAGT 2741

ORIGIN

Query Match		64.4%;	Score 4381.6;	DB 6;	Length 4653;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 4384;		Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	702	CGCAGGAAATCCTGGGTGGCACTGGGCGAAGCCAGAGCCCTTGCACAGCTTGTG	761		
DB	79	CCCCAGGAATCCTGGGTGGCACTGGGCGAAGCCAGAGCCCTTGCACAGCTTGTG	138		
QY	762	GAGGCGGTGAGGACCTGGCCAGAGCTCTGGGCGCTGGCAGCTGGTGGAGCTTCG	821		
DB	139	GAGGCGGTGAGGACCTGGCCAGAGCTCTGGGCGCTGGCAGCTGGTGGAGCTTCG	198		
QY	822	GCACTGTGACAGACCCCGAGGACACAGCGCCCTCTGGAGTTGCTGTCAAGAGCCCTC	881		
DB	199	GCACTGTGACAGACCCCGAGGACACAGCGCCCTCTGGAGTTGCTGTCAAGAGCCCTC	258		
QY	882	TGCAGTGTACAGGGAACCTAGCAGCACAGTGGGCGCCCTCTCAACTGGTACAGGCTAGT	941		
DB	259	TGCAGTGTACAGGGAACCTAGCAGCACAGTGGGCGCCCTCTCAACTGGTACAGGCTAGT	318		
QY	942	GACCTGATGAGAGCTGGTGGGAGAGCGAGATCCGCGCTGCGCAGAGAGCTGAGC	1001		
DB	319	GACCTGATGAGAGCTGGTGGGAGAGCGAGATCCGCGCTGCGCAGAGAGCTGAGC	378		
QY	1002	CCGCGCTCTCGAGCTGATTTGAGCCCTGAGACGACCCGCTGCTCCGCTGCTCTGG	1061		
DB	379	CCGCGCTCTCGAGCTGATTTGAGCCCTGAGACGACCCGCTGCTCCGCTGCTCTGG	438		
QY	1062	AGAGCGCTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGACACAGATACACCTTTTACC	1121		
DB	439	AGAGCGCTGAAGCTCTGATCCTCGGGAAGCTACTCTTTGACACAGATACACCTTTTACC	498		
QY	1122	CGGAGCTCATGCGCCAGGTGAACCGGACCTTGCAGGAGCTCACCTCTGAGGAGTGC	1181		
DB	499	CGGAGCTCATGCGCCAGGTGAACCGGACCTTGCAGGAGCTCACCTCTGAGGAGTGC	558		
QY	1182	CGGAGGTGTGGAGATGCTGGAGCCCGGATCTTCACTTTCATGAAAGAGTTCGAAT	1241		
DB	559	CGGAGGTGTGGAGATGCTGGAGCCCGGATCTTCACTTTCATGAAAGAGTTCGAAT	618		
QY	1242	GTGGCCATGCTGCAGCGCTCTCTGAGATGAGGATGAAGGAAGGAGCCAGACCT	1301		
DB	619	GTGGCCATGCTGCAGCGCTCTCTGAGATGAGGATGAAGGAAGGAGCCAGACCT	678		
QY	1302	GGAGCGCGGACCAATGAGGCGCTGCGATCTTTCTGAGCCCTTGGAGCGGTGGCTAC	1361		
DB	679	GGAGCGCGGACCAATGAGGCGCTGCGATCTTTCTGAGCCCTTGGAGCGGTGGCTAC	738		
QY	1362	AGCTGGAGGACGACAGCTGTGGGCGACCTGGTGGGCGACCTGGGCGGAGTACG	1421		
DB	739	AGCTGGAGGACGACAGCTGTGGGCGACCTGGTGGGCGACCTGGGCGGAGTACG	798		
QY	1422	GAGTGCCTGTCTTGGACAACTGGAGCGGACCCCTCAGAGGAGCCCTGGTGTGGGG	1481		
DB	799	GAGTGCCTGTCTTGGACAACTGGAGCGGACCCCTCAGAGGAGCCCTGGTGTGGGG	858		
QY	1482	GCCCTGCAATCTGTCGCGGAAACATGATTTGGGCGCGCTGCTTCTTGGGACCTGAG	1541		
DB	859	GCCCTGCAATCTGTCGCGGAAACATGATTTGGGCGCGCTGCTTCTTGGGACCTGAG	918		
QY	1542	GACTCTTCAGACCCACAGACACCCACACCTGGGCGCCGCGGACCTGGGCGATC	1601		
DB	919	GACTCTTCAGACCCACAGACACCCACACCTGGGCGCCGCGGACCTGGGCGATC	978		

Db	2059	CGGGGCTCAGCCTGGAGCTTACACAGGGCCACATCACCGCCTTCTGTGGCCACACAGGG	2118
Qy	2742	GCCGGCAAGACACACCCCTGTCTCATCTTGAGTGGCCCTTCCACCCAGTGTGGCTCT	2801
Db	2119	GCCGGCAAGACACACCCCTGTCTCATCTTGAGTGGCCCTTCCACCCAGTGTGGCTCT	2178
Qy	2802	GCCTTCATCTGGGCCACAGAGCTCCGCTCCAGATGGCCGCATCCGGCCCACTGGGC	2861
Db	2179	GCCTTCATCTGGGCCACAGAGCTCCGCTCCAGATGGCCGCATCCGGCCCACTGGGC	2238
Qy	2862	GTCTGTCTCAGTACACAGCTGTCTTGTGACATGTGACCCGTGGACAGACAGCTGTGTTT	2921
Db	2239	GTCTGTCTCAGTACACAGCTGTCTTGTGACATGTGACCCGTGGACAGACAGCTGTGTTT	2298
Qy	2922	TATGGGCGGTGAAGGCTGTAGTGGCCGTGTGTAGTGGCCGCCAGAGAGACCGTCTGCTG	2981
Db	2299	TATGGGCGGTGAAGGCTGTAGTGGCCGTGTGTAGTGGCCGCCAGAGAGACCGTCTGCTG	2358
Qy	2982	CAGGATGTGGGGTGGTCTCCAGCAGAGTGTGCAGACTCGCACCTCTCTGTGGGATG	3041
Db	2359	CAGGATGTGGGGTGGTCTCCAGCAGAGTGTGCAGACTCGCACCTCTCTGTGGGATG	2418
Qy	3042	CAACGGAAAGCTGTCCGTGGCCATTGCTTGTGGGGCGGCTCCCAAGTTGTTATCTTGGAC	3101
Db	2419	CAACGGAAAGCTGTCCGTGGCCATTGCTTGTGGGGCGGCTCCCAAGTTGTTATCTTGGAC	2478
Qy	3102	GAGCTACGGCTGGCGTGGATCTGTCTTCCCGCGCGGTATTTGGAGTGTGCTGCTCAAA	3161
Db	2479	GAGCTACGGCTGGCGTGGATCTGTCTTCCCGCGCGGTATTTGGAGTGTGCTGCTCAAA	2538
Qy	3162	TACGGAGAGGTGCAGCTGATCTCTCCACCACCACTGATGAGGAGAGCTGCTG	3221
Db	2539	TACGGAGAGGTGCAGCTGATCTCTCCACCACCACTGATGAGGAGAGCTGCTG	2598
Qy	3222	GGAGACCGTGTGGTGTGGAGGTGGCGGCTTGTGCTGTGTGGCTCCCACTCTTC	3281
Db	2599	GGAGACCGTGTGGTGTGGAGGTGGCGGCTTGTGCTGTGTGGCTCCCACTCTTC	2658
Qy	3282	CTGCGCGCTCACCTGGGCTCCGCTACTACTGACGCTGTGTGAAGGCCCGCTTGCCTG	3341
Db	2659	CTGCGCGCTCACCTGGGCTCCGCTACTACTGACGCTGTGTGAAGGCCCGCTTGCCTG	2718
Qy	3342	ACCACCAATGAGAGGCTGACATGATGAGGGAGTGTGACACCCAGGAGGAAAG	3401
Db	2719	ACCACCAATGAGAGGCTGACATGATGAGGGAGTGTGACACCCAGGAGGAAAG	2778
Qy	3402	AAGAAATGGCAGCCAGGCGACAGAGTCCGCACTCTCAGCTGTGGCCCTGGTACAGCAC	3461
Db	2779	AAGAAATGGCAGCCAGGCGACAGAGTCCGCACTCTCAGCTGTGGCCCTGGTACAGCAC	2838
Qy	3462	TGGGTGCCGGGCA CGGCTGGTGGAGAGCTGCCACACAGAGTGTGTGGTGTGCC	3521
Db	2839	TGGGTGCCGGGCA CGGCTGGTGGAGAGCTGCCACACAGAGTGTGTGGTGTGCC	2898
Qy	3522	TACAGGGTGGCCATGACGGACGCTTCCGACACTCTTCCGAGAGCTAGACACGCGGCTG	3581
Db	2899	TACAGGGTGGCCATGACGGACGCTTCCGACACTCTTCCGAGAGCTAGACACGCGGCTG	2958
Qy	3582	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCACTCTTCCGAGAGCTAGAC	3641
Db	2959	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCACTCTTCCGAGAGCTAGAC	3018
Qy	3642	AAGGTGTGAGGAGTGTGCTGGGACACAGATATGAGAGTGGCAGCTGCGGGCAGCAC	3701
Db	3019	AAGGTGTGAGGAGTGTGCTGGGACACAGATATGAGAGTGGCAGCTGCGGGCAGCAC	3078
Qy	3702	CTATGACAGGCAATGCTGGCTAGAGCTAACCTTGGGCTCAAGATGCCGCCACAGGAG	3761
Db	3079	CTATGACAGGCAATGCTGGCTAGAGCTAACCTTGGGCTCAAGATGCCGCCACAGGAG	3138
Qy	3762	ACAGCGCTGGAGAACGGGGACAGCTGGGTTCAGCCGCCAGAGACTGACACAGGGCTTGGG	3821

Db	3139	ACAGCGCTGGAGAACGGGGAAACGAGCTGGGTTCAGCCCGCCAGAGACTGACACAGGGCTCTGGG	3198
Qy	3822	CCAGACGCGGTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGCAGCAGCTTCCAGGCCCTG	3881
Db	3199	CCAGACGCGGTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGCAGCAGCTTCCAGGCCCTG	3258
Qy	3882	CTTCTCAAGCGCTTCTGTCTTGGCCCGCCGCGCAGCGCGGGCCCTGTTCGCGCCAGATCGTG	3941
Db	3259	CTTCTCAAGCGCTTCTGTCTTGGCCCGCCGCGCAGCGCGGGCCCTGTTCGCGCCAGATCGTG	3318
Qy	3942	CTGCTGCCCTCTTTTGTGGGCTTGGCCCTGTGTTCAGCCTCATCTGTCCTCTTTTCGGG	4001
Db	3319	CTGCTGCCCTCTTTTGTGGGCTTGGCCCTGTGTTCAGCCTCATCTGTCCTCTTTTCGGG	3378
Qy	4002	CACCTACCGGCTCTGGGCTCAGTCCCAATGTAGGTGCTCAGGTGCTCTTCTTCAGT	4061
Db	3379	CACCTACCGGCTCTGGGCTCAGTCCCAATGTAGGTGCTCAGGTGCTCTTCTTCAGT	3438
Qy	4062	GAGGACGCGCCACAGGGACCTCTGAGCGTCCCGGCTGCTCGAGGCGCTGTGTCAGGAGGCA	4121
Db	3439	GAGGACGCGCCACAGGGACCTCTGAGCGTCCCGGCTGCTCGAGGCGCTGTGTCAGGAGGCA	3498
Qy	4122	GGACTGGAGAGAGCCCCAGTGCAGCATAGTCCCAAGTTCCTCGGCAACAGAAATTCCT	4181
Db	3499	GGACTGGAGAGAGCCCCAGTGCAGCATAGTCCCAAGTTCCTCGGCAACAGAAATTCCT	3558
Qy	4182	GCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGACCCAGAGTCTCCATCCCCAGCC	4241
Db	3559	GCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGACCCAGAGTCTCCATCCCCAGCC	3618
Qy	4242	TGCGAGTGTAGCCAGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGAGCTGGT	4301
Db	3619	TGCGAGTGTAGCCAGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGAGCTGGT	3678
Qy	4302	GGTCCCGCTCCGCGCCAGGACAGTACCGGCTCTGGGAAAGTGTTCAGAACTTGACAGGC	4361
Db	3679	GGTCCCGCTCCGCGCCAGGACAGTACCGGCTCTGGGAAAGTGTTCAGAACTTGACAGGC	3738
Qy	4362	CGGAACTGTCTGACTTCTGTGTCAAGACCTACCCGCGCTGTGTGCGCGCGCTGGAAG	4421
Db	3739	CGGAACTGTCTGACTTCTGTGTCAAGACCTACCCGCGCTGTGTGCGCGCGCTGGAAG	3798
Qy	4422	ACTAAGAGTGGGTGAATGAGTACGAGGCTTCTCGCTGGGGGCGGAGACCCA	4481
Db	3799	ACTAAGAGTGGGTGAATGAGTACGAGGCTTCTCGCTGGGGGCGGAGACCCA	3858
Qy	4482	GGCTTCCCTCGGGCCAAAGTTGGCCCGCTCAGTGGAGAGTTGTGGCGCTGTGAGT	4541
Db	3859	GGCTTCCCTCGGGCCAAAGTTGGCCCGCTCAGTGGAGAGTTGTGGCGCTGTGAGT	3918
Qy	4542	CCCCTGCTGGGGGGCCCTCGACCGTGTCTGAAAACCTCACAGCCTGGGCTCACAGC	4601
Db	3919	CCCCTGCTGGGGGGCCCTCGACCGTGTCTGAAAACCTCACAGCCTGGGCTCACAGC	3978
Qy	4602	CTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAAACAAAGGCTGGGCACTCCATGGTG	4661
Db	3979	CTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAAACAAAGGCTGGGCACTCCATGGTG	4038
Qy	4662	GCCTTTTCAACCGAGCCAGCAACGCAATCTCCGCTGCTCACCTGCCCCCAGGCGCGGCC	4721
Db	4039	GCCTTTTCAACCGAGCCAGCAACGCAATCTCCGCTGCTCACCTGCCCCCAGGCGCGGCC	4098
Qy	4722	CGCCACGCCCAAGCATACACACACTCAACACCCCTTGNACCTCACAGGAGCAGCTG	4781
Db	4099	CGCCACGCCCAAGCATACACACACTCAACACCCCTTGNACCTCACAGGAGCAGCTG	4158
Qy	4782	TCTGAGGCTGCACTGATGGCTCTCGGTGGAGCTCTCTGCTTCCATCTGTGTGGTCTTT	4841
Db	4159	TCTGAGGCTGCACTGATGGCTCTCGGTGGAGCTCTCTGCTTCCATCTGTGTGGTCTTT	4218
Qy	4842	GCCATGTCTTTGTCCCGGCCAGCTTCACTTTGTCTCATTTGAGGAGCGAGTCAACCGA	4901
Db	4219	GCCATGTCTTTGTCCCGGCCAGCTTCACTTTGTCTCATTTGAGGAGCGAGTCAACCGA	4278

```
QY 4902 GCCAAGCACCTCAGCTCA TGGGGGGCTGTCCCCACCCCTCTACTGGCTTGCGCACTTT 4961
DB 4279 GCCAAGCACCTCAGCTCA TATGGGGGGCTGTCCCCACCCCTCTACTGGCTTGCGCACTTT 4338
QY 4962 CTCCTGGGACATGTAACTACTTGGTGGCAGCATGCTGCTGCTCATCTTTCTGGCC 5021
DB 4339 CTCCTGGGACATGTAACTACTTGGTGGCAGCATGCTGCTGCTCATCTTTCTGGCC 4398
QY 5022 TTCACGACAGGCGCATATGTGCGCCCTGCGCAACCTGCGCTCTCTCTGCTGTGCTACTA 5081
DB 4399 TTCACGACAGGCGCATATGTGCGCCCTGCGCAACCTGCGCTCTCTCTGCTGTGCTACTA 4458
QY 5082 CTGTATGG 5089
DB 4459 CTGTATGG 4466

RESULT 13
AB097814
LOCUS
DEFINITION
Rattus norvegicus ABCA7 mRNA for ATP-binding cassette transporter
sub-family A member 7, complete cds.
ACCESSION
AB097814
VERSION
AB097814.1 GI:33411635
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Sasaki, M., Nada, S. and Yamaguchi, A.
Cloning of rat ABCA7
Unpublished
2 (bases 1 to 6513)
Sasaki, M., Nada, S. and Yamaguchi, A.
Direct Submission
Submitted (09-DEC-2002) Mari Sasaki, I.S.I.R., Osaka University;
Mihogaoka 8-1, Ibarakishi, Osaka 567-0047, Japan
(E-mail: sasaki37@osaken.osaka-u.ac.jp, Tel: 81-6-6879-8546,
Fax: 81-6-6879-8549)
FEATURES
source
location/Qualifiers
1 . 6513
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wister"
/db_xref="taxon:10116"
/tissue_type="platelet"
/dev_stage="adult"
1 . 6513
/gene="ABCA7"
1 . 6513
/gene="ABCA7"
/codon_start=1
/product="ATP-binding cassette transporter sub-family A
member 7"
/protein_id="BAC81426.1"
/db_xref="GI:33411636"
translation="MAFCTQLMLLLKRNYYRRQPIQLVVELLWPLFLFFILVVRH
SHPLEHECHFPNKPPLPSAGTVPMQLGVNVCNNSCFQPTPGKPGVLSNFKDSLI
SRLLADHATVLGCHSTQDMLAALGKLI PVLRVAGSGAWPOESNOPAKGVSVELLEKI
LORASLETVLGQAODSMKFSDATRTVAQELLTLPSELVRLALLRRPRGSGSLELIS
EALCSTKGPSPGSLNWEANQINPEMPGPELAPTL PDSSLSPACSEFVGCALDDHPV
SRLLWNTVMQLKLDDVEGTGQQQT PKGQKLEAIRDFLDPSGRYNNQBEAHDMGR
LAETLGQLECVLKDLEAVPSEALNSRALELLGERRLWAGI VFLSPHPHLDSEPP
SPTTTPGHRLVKIRMDIDVTRTNKIRDFKWDPPSADPLMDLRYVMGFGVYLQDILL
EQAAVRVLSGRDSRAGLYLOQMPHCYVDVFLVLSPLFLFLTAWIYSVALTVKA
VVRKETRLRMTAMGLSRVAILWGLFSLCGLGEPFLVSALLVLVLKGLNLPYSHPV
VVFLLFAAFATVAQSPFLSAFFSRANLAAACGLAYFALYLPVLCVAVWRRLPLG
GLLAVLSLPVAFGFGESLALBEEQGDGAQHNLGTCPAEDVFLSLAQVSAFLLLDPAV
IYGLAWLYLEAVCPQGVGIPPEPNFPFRSRYWCGPGPPKSSVLAPAPQDPKVLVEEPP
```

ORIGIN

```
Query Match 57.4%; Score 3902.6; DB 10; Length 6513;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 4982; Conservative 0; Mismatches 1429; Indels 132; Gaps 9;
QY 210 ATGGCTTCTTGACACAGCTGCTGCTGTCTGGAAGAAATTCATGTATCGCGGAGA 269
DB 1 ATGGCTTCTTGACACAGCTGCTGCTGTCTGGAAGAAATTCATGTATCGCGGAGA 60
QY 270 CAGCGGTCCAGCTCCTGCTGAATGCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCT 329
DB 61 CAACCGATCCAACTAGTAGTGAGTACTGTGGGCCCTCTCTCTCTCTCTCTCTCT 120
QY 330 GCTGTGTCCTCCACTCCACCCCTGAGCAGCAGCATGATGCTCTCTCTCTCTCTCTCT 389
DB 121 GCGGTCTCGTCTACTACACCCCTGAGCAGCAGCATGATGCTCTCTCTCTCTCTCT 180
QY 390 CTGCCATCGCGGGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449
DB 181 CTACCATCAGCGGGCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 450 TGCTTTCCGAGCTGACACCGGGGAGAGAGCCCGGGGCGCTGAGCAA CTTCACAGCTCC 509
DB 241 TGCTTTCCAAACACCCAAACCGCTGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTG 300
QY 510 CTGGTCTCCCGGCTGCTAGCCGATGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
DB 301 TTGATCTCGAGGCTCTCTGCGCATGCTCACAACAGTGTCTGGGGGGCCACAGACCCAG 360
QY 570 ACCTGCTGCTGCTAGGGAAGCTGATCGCCAGCTGAGGGCTGCGCAGCAGCAGCCCGC 629
DB 361 ATGCTGCTGCTGCTGCGGAAGCTGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 630 CCTCAACCAACCAAGCAGCTCTCCA CTGGAA CCA CCA CCA TGTCTGATGTCTGCG 689
DB 421 CCACAGGAGAGTAACCAACCGCTAAGCAAGGCTCAGTGA -----CTGAGCTCTTG 471
QY 690 ACCTCACTGCTGCGCAGGATCCCTGGGGTTGGCACTGCTGGGCGAAGCCAGGAGCCCT 749
DB 472 GAGAAGATCTCTGCAACCGGCTATCCCTGAGACCCGTGCTGGGTCAAGCCAGATTCAT 531
QY 750 CACAGCTTGTGGAGGCGCTGAGGA CTGCGCCAGGAGCTCTCTGCGCGCTGCGCAGCTG 809
DB 532 AGAAGTCTCAGATGCTACCAAGAACTGTTGCTCCAGGAGCTCTCTGACCTACCCAGCT 591
QY 810 GTGAGCTTCTGGGCACTGCTGTCAGAGACCCCGAGGAGCACAGCGGCCCTCTGGAGT 869
DB 592 GTGGAACCTCCGGGCTTGTCTGCGGAGGCCCGAGGGT CAGCTGCTGCTCTCTGAGCTGAT 651
```

Qy	870.	TCAGAGGCCCTCTGACGTGTACGGGACCTAGCAGACACAGTGGGCCCTCCCTCAACTGG	929
Db	652	TGGAGGCCCTCTGCAGTACCAAGGGGCCAGCAGTCCAAGGGGTCTGTCCCTCAACTGG	711
Qy	930	TACGAGGCTAGTGACTGATGGAGCTGGTGGGCGAGGCGAGACAGAAATCCGCCCTGCCAGAC	989
Db	712	TACGAGCCCAACAGATCAACGAATTCATGGGCTGAGCTGGCGCCACCCCTGCCTGAC	771
Qy	990	AGCAGCTGAGCCCCGCTGCTGGAGCTGATTTGGAGCCCTGGACACCCACCCGCTGTCC	1049
Db	772	AGCAGTCTACGCCCGCTGCTGTAGTTCTGTGGGGCACTGGATGACCACTGTGTGTCT	831
Qy	1050	CGCCTGCTTGGAGACGCTTAAGCCTCTGATCCTCGGAGCTACTCTTTTGACCAAGAT	1109
Db	832	CGGCTGCTTGGAGACGCTTAAGCCATTGATCCTCGGAAAAATCTCTTTTGACCACTGAC	891
Qy	1110	ACACCTTTTACCGGAGCTCATGGCCAGGTGAAACCGGACCTTCGAGGAGCTCACCCCTG	1169
Db	892	ACAACTTCACTAGGAAACTCATGGCTCAGGTGAAACAGACCTTTGAGGAGCTGGCTCTG	951
Qy	1170	CTGAGGATGTCCGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTTCACCTTCATGAAC	1229
Db	952	CTGAGGACCTACACGAACTCTGGGAGTACTGGGACCCAGATCTTCACTTCAATGAAT	1011
Qy	1230	GACAGTTCGAATGTGGCCATGTGCACGGCTCTCTGCAGATGCAGGATGAAGAAAGG	1289
Db	1012	GACAGTACAAACGTGGCCATGCTTCAGAAACTGCTGGATGTGGAGGCGACAGGCTGGCAG	1071
Qy	1290	CAGCCAGACTGGAGGCGGGACCAATCGAGGCGCTGCGATCTTCTGGGACCCCTGGG	1349
Db	1072	CAGCAGACACCCAAAGGCCAAGACGTTGGAGCTATCAGAGACTTCTTGGATCCTTAGT	1131
Qy	1350	AGCGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGGCACCTGTGTGGGACGCTG	1409
Db	1132	AGGGTCGATACAACTGGCAGGAGGCTCATGCAGACATGGGCGCCTGGCAGAAATACTG	1191
Qy	1410	GGCGAGTGACGAGTGCCTGTCTTGGACAGCTGGAGGGGCACTCCTCAGAGGCGCC	1469
Db	1192	GGCCAAATCTTGGAGTGTGTCTCCCTGGACAGCTGGAGCTGTGCCCTCAGAGGAAGCT	1251
Qy	1470	CTGTGTTCGGGGCCCTGCAACTGCTCGCGAACTCGATTTCTGGGCGCGCTGCTCTTC	1529
Db	1252	CTTGTGTTCGTGCTTGGAGTCTGTGGGTGAGCGCGCTCTGGGAGGACATCTCTTC	1311
Qy	1530	TTGGGACCTGAGGACTCTTCAGACCCCAAGAGACCCCAACCCCA---GACCTGGGCCCC	1586
Db	1312	CTGAGCCAGAGCACCTCTGGAATCATCTCAACCGGCATCTCAACCAACAGGGTCTCT	1371
Qy	1587	GGCCAGTGGCATCAAAATCCGATGGAATTCAGCTGTGTACGAGGACCAATATAGATC	1646
Db	1372	GGCCACTACAGTCAAGATTCGAATGGATATTGATGATGTCAACAGGACCAATAAGATC	1431
Qy	1647	AGGGACAGGTTTGGGACCTTGGGCCAGCGCGGACCCCTGACCGGACCTTCGCTACGCTG	1706
Db	1432	AGGGACAAGTTTGGGACCCAGTTCCTGACAGATCCCCTCATGAGACCTTCGGTATGTG	1491
Qy	1707	TGGGGCGGCTTCGTGTACTGTCAAGACCTCGGTGGAGCGTGCAGCCGTTCGGCTCAGC	1766
Db	1492	TGGGGCGGCTTCGTGTACTGTCAAGACCTGCTGGAGCAGGAGCTGTGTGCGCTGCTGAGT	1551
Qy	1767	GGGCCAAACCCCGGGCGGCTCTACTCGACAGATGCCCTATCCGTCTGATGTGGAC	1826
Db	1552	GGCAGGACTCCAGCAGGCGCTCTATTGGACAGATGCCACACCCCTCTCTACGTGGAC	1611
Qy	1827	GACGTGTTCTGCTGTGCTGAGCCGCTCGCTGCGCTCTTCTCTGAGCTGGGCTGGATC	1886
Db	1612	GATGTGTTCTGCGGGTCTGAGCGGTCTTGCCCTCTGTCTCTGATCTCTGCGCTGGATC	1671
Qy	1887	TACTCCGTGACATGACAGTGAAGGCCCTGGTTCGGGAGAAAGGAGACCGCGCTCGGGAC	1946
Db	1672	TACTCGTGGCGCTCACTGTGAAGGCCCTGGTTCGGCGAGAAAGAGACACGGCTCGAGAA	1731

QY	1947	ACCAATGCGCGCATATGGGGCTCTACGCGCGCGGTGCTCTGGCTAGGCTGGTTCTCTCAGCTGC	2006
DB	1732	ACCAATGGCTGCGATGGGGTTTACGCGGTGCTGTGCTCTGGCTTGGTTGGTTTCTCTCAGCTGC	1791
QY	2007	CTCGGGCCCTTCCGTCTCAGGCGCGCGCTGCTGGTTCTGGTGCTCAAGCTGGGGGACATC	2066
DB	1792	CTGGGAACCTTCTTGGTCAAGCGCTGCAATGTGTGGTGTAGTGCTTAAGCTAGGGAAATC	1851
QY	2067	CTCCCTTACAGCACCGCGCGTGGTCTTCCCTGTTCTTGGCAGCCTTCGCGGTGGCCACG	2126
DB	1852	CTTCTCTTACAGCACCGGTGTAGTCTTCTTCTTCTTGGCGGCTTGGGTGGCCACC	1911
QY	2127	GTGACCCAGAGCTTCTGCTCAGCGCCTTCTTCTCCGCGCCAACTGGCTGGGCGCTGC	2186
DB	1912	GTGCTCAGAGTTTTCTGCTCAGTGCCCTTCTTCTCCGCGCCAACTGGCAGCAGCCTGC	1971
QY	2187	GGCGGCTGGGCTACTTCTCCCTACTACCTGCCCTACGTGCTGTGTGGCTTGGCGGAC	2246
DB	1972	GGGGGCTTAGGCTACTTTCGGGCTCTACTCCCTCATGTGTGTGTGCTGTGCGCTGGCGTGA	2031
QY	2247	CGGCTGCCGCGGGTGCGCGTGGCGGAGCCTGTCTGCTCGCCGTGGCTTTCGGCTTC	2306
DB	2032	CGCTCTGCTCTGGCGGACTTCTTAGCTGTGAGCTGTCTGTCTCCCGTGGCCCTTGGCTTT	2091
QY	2307	GGCTGCGAGACCTGGCTCTGCTGGAGGAGCAGGGCGAGGGCGCGAGTGGCAACGCTG	2366
DB	2092	GGATGCGAGAGCCTGGCTCTACTTGGAGGAGCAGGGGACGCGGCTCAGTGGCACAAATTTG	2151
QY	2367	GGCACCCGGCTACGGCAGAGCTTTCAGCCTGGCCAGGTCTCTGGCCTTCTGCTGCTG	2426
DB	2152	GGCACAGGCCCGCGGAGGATGTCTTCAGCCTGGCACAGGTGTCTGCTTCTCTGTTGCTT	2211
QY	2427	GACGCGGCTCTACGCGCTCGCCACCTGTGTACTGTGAAGCTGTGTGCCCGCAGGCCAGTAC	2486
DB	2212	GATGCTGTATCTACGGCCCTTGCCCTCTGGTACTTAGAGGCTGTATGCCCAGGCCAGTAT	2271
QY	2487	GGGATCCTCGAACCATGGAAATTTCTTTTTCGAGGAGTCTACTGTGTGCGGACCTCGGCC	2546
DB	2272	GGAACTCCTGAACCATGGAAATTTCCCTTTTCGGAGGAGCTACTGGTGTGGCCCTTGACCT	2331
QY	2547	CCGAAGAGTCCAGCCCTTTCGCCACCCCGCTGGACCCCAAGGTGCTGTTAGAGAGGCA	2606
DB	2332	CCCAAGAGTCTGTCTTGGCCCTTGCCCCACAGATCCCAGGTTCTGTGTGGAGAGGCCA	2391
QY	2607	CCGCCCGGCTGAGTCTCTGGCGTATCCGTTTCGAGCCTGGAGAGCGCTTTCCTGTGAAGC	2666
DB	2392	CCACTGGTCTGGTTCTCTGTGTCTCATTTGAGGCTGGAAGCAATTTTCTGGTCTCT	2451
QY	2667	CCGAGCCAGCCCTGCGGGGCTCAGCCTGGACTTTTACGAGGCGCACATCAGCCCTTC	2726
DB	2452	CCAAGCCAGCCCTGCGAGGGCTCAACCTTTGACTTCTTATGAAGGCCACATCACTGCCTTT	2511
QY	2727	CTGGGCCACAACGGGGCCGGAAGACACACCTGTCCATCTTGTAGTGGCTTCTTCCCA	2786
DB	2512	TTGGGTCACAACGGGGCTGCAAGACAAACACACTGTCCCATCTTGTAGTGGTCTCTTCCCA	2571
QY	2787	CCAGTGTGTGCTGTGCTTTCACTCTGGGCAAGAGCTCGCTCCAGCATGGCGCCATC	2846
DB	2572	CCCAAGTAGTGGCTCGGCTCTCATCTCTGGGCGATGACGTGAACCAACATGGCAGCCATC	2631
QY	2847	CGGCCCAACCTGGGCGTCTGTCTCAGTACAAAGTGTCTTTGACATGCTGACCGTGGAC	2906
DB	2632	CGGCCCCACCTGGGCATCTGCCACAGTACAAATGTGTGTTTGAATGTCTGACOGTGA	2691
QY	2907	GAGCACGTCTGTCTATGGGCGGCTGAAGGCTAGTCCGCTGTAGTGGGCCCGGAG	2966
DB	2692	GAAACGTTTTGGTTCTATGGCCGTTTGAAGGGCGTGAAGTCAAGCGGCTATAGACTCCGAG	2751
QY	2967	CAGCACGTCTGTCTCAGGATGTGGGCTGTCTTCCAAGCAGAGTGTGCAGACTCTCGCCAC	3026
DB	2752	CAGGAACAATTTGATTGGGATGTGGGCTCATCCCTAAGCGGACACACAGACGCGTAC	2811
QY	3027	CTCTCTGGTGGGATGCAACGGAAAGTGTCTCGTGGCCATTCCTTTGTGGCGGCTCCCAA	3086

QY 5148 GCCTATGTGTGCTCACATTAACCTCTTTATTTGGCATCAATGGAAGCATGGCCACC 5207
DB 5014 GCCTATGTGTGTCTCACTGATCAACCTCTTTATTTGGCATCAATGGAAGCATGGCCACC 5073
QY 5208 TTTGTGCTTGTGCTCTCTCTGATCAGAAGCTGAGGAGGTGAGCGGATCTTTGAAACAG 5267
DB 5074 TTGTGCTGTGAACTGTCTTTTGTGATCAGAACTTCAAGAGTGTGAGCGGATCTTGAAGCAA 5133
QY 5268 GTCTTCTCTTATCTTCCCTCACTTGTCTTGGCGCGGGGCTCATGTGATGATGGTGGGAAAC 5327
DB 5134 GTGTTTCTTATCTTCCCTCACTTGTCTTGGCGCGGGGCTCATGTGATGATGGTGGGAAAC 5193
QY 5328 CAGGCCATGTGCTATGCTTTGAGCGCTTGGAGAGCAGGAGTTCAGTCAACCTCTGCGC 5387
DB 5194 CAGGCCATGTGCTATGCTTTGAGCGCTTGGAGAGCAGGAGTTCAGTCAACCTCTGCGC 5253
QY 5388 TGGAGGTGTGCGCAAGAACCTCTTGGCCATGCTGATCAGAGGCGCCCTCTTCTTCTC 5447
DB 5254 TGGAGCATCAATGGCAAGAACCTCTTGGCCATGCTGAGCGCTTGGAGAGCAGGAGTTCAGTCAACCTCTTCTC 5313
QY 5448 TTCACTACTGTGTGAGCAGCAGAACCTCTTGGCCATGCTGATCAGAGGCGCCCTCTTCTC 5507
DB 5314 ATCACACTCTCTGCTCCAACTGCAACCGCTCTCTGCGCAACCAAAATCAAGACTGCGC 5373
QY 5508 CCACTCTGTGAGAGGAGGAGGAGTGTAGCCCGTGAACCGGAGCGGCTGCTCAAGGA 5567
DB 5374 CCGCCCTCTGGGAGGAGGAGTGTGCTGAGAGCGCTGAGCGGCTGCAAGGGA 5433
QY 5568 GCCACCCAGGGGATGTTGGTGTCTGAGGAACTTGACCAAGGTATACCGTGGGAGAGG 5627
DB 5434 GCCACCCAGGGGATGTTGGTGTCTGAGGAACTTGACCAAGGTATACCGTGGGAGAGG 5493
QY 5628 ATGCCAGCTGTGACCGCTTGTGCTGGGATTCCTCTGCTGAGTGTGTTGGGCTGCTG 5687
DB 5494 AGCCAGCAGTGTGATCACCTGTGCTTGGGATTCCTCTGCTGAGTGTGTTGGGCTGCTG 5553
QY 5688 GGTGTGATGAGGAGGAGAGAGTGTGCTGCTGAGTGTGAGCGGAGGAGTGTGCTGCTG 5747
DB 5554 GGTGTGATGAGGAGGAGAGAGTGTGCTGCTGAGTGTGAGCGGAGGAGTGTGCTGCTG 5613
QY 5748 AGCAGGGCGAGGCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5807
DB 5614 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5673
QY 5808 CTGAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5867
DB 5674 CGCAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5733
QY 5868 CACTGAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5927
DB 5734 CACTGAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5793
QY 5928 GGTCTGAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5987
DB 5794 CTCTGAGCATGGGATCTGCTTCAATCCGATGCCATCTTTGAGTGTGAGCGGCGCGAG 5853
QY 5988 AGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6047
DB 5854 AGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5913
QY 6048 GTGTTTCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6107
DB 5914 GTCTTCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5973
QY 6108 AGCTTTTGTGCGGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6167
DB 5974 AACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6033
QY 6168 GAGTGTGAAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6227
DB 6034 GAGTGTGAAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6093

QY 6228 GGCGAGCCCGCAACATCTCAAGGCGAGATTCGCGGCGGCTCACACATGACCCCTCGCGTG 6287
DB 6094 GGAAGCGCTCAGCATCTCAAGAGCAGATTCGCGGCGGCTGGCCACACATGACTCTCAGGTC 6153
QY 6288 CCGCGCGCAAGGTCACAGCGCGGCGGCTTCTGTCGCGCGGAGTTCCTCTGCGTGGGTCGAG 6347
DB 6154 CCACCGGACCGCGGAAACCGCGGATAGCTTCTATTTGTGACACATTTCTCTGATGCTGAG 6213
QY 6348 CTGCGGAGGAGCAGATGAGCGCGCTGCGCTTCCAGCTGCGCGCGGAGGCGCTGGCC 6407
DB 6214 CTTTCGCGAGGTGATGCGCAGCGCTTCTGCGCTTTCAGCTGCGCGCTGGGCGGATGACG 6273
QY 6408 CTGCGCGCGCTCTTTCGAGAGCTGCGCGTGCAGCGCGCAGAGCACGCGCTGGAGGACTTT 6467
DB 6274 CTGCGCAGAGTGTTCAGGAGCTGCTGCGCAGGCAAGGCCCAACGCTGTGAGGACTTC 6333
QY 6468 TCGTGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTCTTCCAAAGGACCGAGGAG 6527
DB 6334 TCTGTGAGCCAGACCACTTTTGGAGGAGTGTCTTATTTCTTCCAAAGGACCGAGGAG 6393
QY 6528 GAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6587
DB 6394 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6450
QY 6588 CAGCAGCCCAAGCGCTGAGCGGCTTCTGATGAGCTTCTGAGGAGGAGGAGGAGGAGGAGG 6647
DB 6451 CAGCATCCCAAGCGGCTGAGCGGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6510
QY 6648 TGA 6650
DB 6511 TGA 6513

RESULT 14
AX644619
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX644619
Sequence 1 from Patent WO0209108.
AX644619
AX644619.1 GI:28610635
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Shutter, J. and Ulias, I.
Atp-binding cassette transporter-like molecules and uses thereof
Patent: WO 0209108-A 1 12-DEC-2002;
Amgen, Inc. (US)

Location/Qualifiers
1..6633
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
1..6504
/note="unnamed protein product"
/protein_id="CAD68131.1"
/gb_xref="GI:28610635"
/translation="MAFCTQLMLLWKNYTRRRQIQLLVLLWPLFLPFLVAVRH
SHPLSHHCCHPNKPLPSAGTVPWLQGLVCNNNSCFQHPPTGEXPGVLSNFKDSLI
SRLADTRVLVGHSSIDMDLALGKLIPLVAVGGARPOESDQPTSGSVTKLEKI
LQASLDPVLGQDQSNRXPDAIRDLAELLTLPSELRLRRRPSKLELVLS
EALCSGTPSSPGSLNRYEANQLNFMGPEVAPALPNLSLSPACSEFVGTUDDHPV
FLLMRRLKPLILGKILFAPDNFTNFKMAQVNTFEELALLDLHLWGLVLPQIFN
SWNDSTNVALQRLDVGQTQQTQPAKKEAIKDFLPSRGYSWREARADMR
LAGILQMECEVSLDKLEAVPSEALVSRLELLBRRRLWAGIVLSPHPLDPSLS
SPALSPHRLFKIRMDIDVTRTNKIRDFMDPSPADPMDLRYVMGPFVYLQDLLE
QAAVRVLGGNSRTGLYLQMPHPCVDDVFLRLVLSPLFLTLAWISVATVKA
VREKTRLMETRMAGLSRAVLWGLFSLCLGFLVSAALLVLKLVLTLPSPHPV
IFLFLAFAVATVAQFLSLSAFSAFLAAACGLAYFALYFLVLCVAVRRLHUGG
LIAASLLSPVAFGFCESLALLEEQDGAQWHLNLTGPAEDVFLSAQVSAFLLDVI
YGLAIWYLEAVCPGQYGIPEPWNFPFRSWCQGPFPKSSVLAPAPQDPKVLVEEPL

Db 3868 CAGGTCTCGTTCTTACGTGAGATGCCCTTGGGACCCCAACCGGATGAAGTGTCTGGAG 3927
Qy 4104 GCGTCTCTCAGGAGCAGACTGGAGAGCCCCAGTGCAGATAGTCTCCACAGG--- 4160
Db 3928 GCTCTGTAGGGGAGGCTGGGCTGCAGGAACCCAGTATGCAGGACAAAGATGCCAGGGA 3987
Qy 4161 -----TTCTGGGCAACCAAGATTTCTCTGAAGTG 4190
Db 3988 TCTAGTGTACACATCTCCCTAGCTTTCAGGTCCTTCAAGTCCCTGAGGTCCCTCTGATGTG 4047
Qy 4191 GCCAAGTCTTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCCAGCCTGCCAGTGT 4250
Db 4048 GCCAGATCTCTGGCCAGTGGCAACTGGACCCAGAGTCTCCATCCCCAGCTTGGCAATGC 4107
Qy 4251 AGCCAGCCCGGTGCCCGGCGCTGTCTGCCGACTGCCCGGCTGCAGCTGGTGGTCCCCCT 4310
Db 4108 AGTCAGCCTGGAGCCCGCGCTGTTCAGATTGCCCGGCTGGAGCTGGGGGTCAACA 4167
Qy 4311 CCGCCCCAGGCAGTGACCCGGCTCTGGGGAAGTGTTCAGAACTGACAGGCCGGAACCTG 4370
Db 4168 CCCCCCAGGCTGTGGCTTGGGAGAGTGGTCCAGAACTCTCACCTGCGCCGAAATGTG 4227
Qy 4371 TCTGACTTCTTGTCAAGACTTACCGCGCTGTGGTGGCCAGGCGCTGAAGACTAAGAG 4430
Db 4228 TCTGACTTCTTGTGAAGACATACCCAGCCTGGTGGCCGAGGCTTAAGACCAAGAG 4287
Qy 4431 TGGGTGAATAGGTACAGGTACCGAGCTTCTCGCTGGGGGCGGAGACCCAGGCGCTGCC 4490
Db 4288 TGGGTGATAGGTACAGATATGGGGCTTCTCCCTGGGAGGCGGAGATCCAGACTTGCCC 4347
Qy 4491 TCGGGCCAGAGTTGGGCGCTCAGTGGAGAGTGTGGGCGCTCTGATGCCCTGCT 4550
Db 4348 ACAGGGCATAGGTGGTCCGCACATTGGCAGAGATTCCGGCACCTGTGAGCCCCAACCT 4407
Qy 4551 GCGGGGCGCTCGACGTGTCTGAAACCTTCAAGCCTGGGCTCAAGCCTGATGCT 4610
Db 4408 GGGAAATGGCTAGACCGTATCTGAAACCTTCACTCAGTGGGCGCTTGGCGCTTATGCT 4467
Qy 4611 CAGGACAGTCAAGATCTGGTTCAACAAGAGGTGCACTCCATGGTGGCCTTCTC 4670
Db 4468 CGGAACGCTCAAGATCTGGTTCAACAAGAGGTGCGATGGCCTGCTGGCCTTGTG 4527
Qy 4671 AACCGAGCAGCAACGCAATCTCGTGTCTACCTGCCCGCCAGGCGCCCGCCACGCC 4730
Db 4528 AACCGAGCAACAAATGGACTCTTACATGCCCCCTTACCACTCTGCTCCAGTCCGCCATGCC 4587
Qy 4731 CACAGCATCACACACTCAACACCCCTTGAACCTTCAAGAGGAGAGTGTCTGAGGCT 4790
Db 4588 CACAGCATCATACACTCAACCACTCTTGAACCTTGAACCTTGAACAGGAGCAGCTATCTGAAGCT 4647
Qy 4791 GCACGTATGGCTCTCGTGGAGCTCTCGTCTCCATCTGTTGTTGCTTTCGCCATGTCC 4850
Db 4648 ACACGTATAGCTCTCTGTGGATGTCTTGTCTCCATCTGTGGTCTTTCGCCATGTCA 4707
Qy 4851 TTTGTCCCGGCGAGCTTCACTCTTCTCTCATTTAGAGGAGCAGTCAACCCGAGCCCAAGCAC 4910
Db 4708 TTTGTCCAGCAGCTTTACCTGTCTCTCATAGAGGAACGATCAACAGAGCCAAAGAT 4767
Qy 4911 CTGACGTCTATGGGGGCTGTCCCCACCTCTACTGGTGTGGCAACTTCTCTGGAC 4970
Db 4768 CTGACGTGGTTCAGCGGCTGCCCAACCCCTTATTTGGTCTTGGCACTTCTCTGGGAC 4827
Qy 4971 ATGTGTAATCTTGGTGGCAGATGCATCGTGGTCTCATCTTCTGGCCTTCCAGCAG 5030
Db 4828 ATGTGTAATCTTGGTGGCAGTGTGCATAGTGGTGTTCATCTTCTAGCCTTTTCAGCAG 4887
Qy 5031 AGGGCATATGGGCGCTTGGCAACCTGCTGCTCTCTGCTGTGTGCTACTACTGTATGSC 5090
Db 4888 AGAGCTATGTGGCCCCAGAGAACCTGCTGCTCTTACTTCTGCTTCTGCTGTATGGG 4947
Qy 5091 TGGTCGATCACACCGCTATGTACCCAGCCTCTTCTTCTCTCGTGGCCAGCAGGCC 5150
Db |||||

Db 4948 TGGTCTATCACACCACTCATGTACCCAGCCTCTTCTTCTTCTCAGTGGCCAGCAGCGCC 5007
Qy 5151 TATGTGGTGTCTCACCTGCATAACTCTTTATTTGGCATCAATGGAGCATGCGCACTTTT 5210
Db 5008 TATGTGGTGTCTCACCTGCATCAACTCTTCTCATTTGGCATCAATAGCAGCATGCGCACCTTC 5067
Qy 5211 GTGTGTGAGCTTCTCTCTGATCAGAACTGTCAGGAGGTGAGCCGGATCTTGAACAGGTC 5270
Db 5068 GTGTGTAGAACTGTCTTTCAGATCAGAACTTGCAGAAAGTGGCCGGATCTCTGAAACAAGTG 5127
Qy 5271 TTCTCTATCTTCCCCCACTTCTGTCTGGCCGGGGGCTCATTTGACATGGTGGCGGAACGAG 5330
Db 5128 TTTCTTATTTTCCCCCACTTTTGGCTTGGCCGAGGGCTCATTTGACATGGTGTTCGGAACGAG 5187
Qy 5331 GCATGGGTGATGCTTTTGGCGCTTGGGAGACAGCGAGTTCAGTCAACCCCTGCGCTGG 5390
Db 5188 GCCATGGCAGATGCTTTTGGCGCTTAGGAGACAGCAATTTTCACTACCCCTACGCTGG 5247
Qy 5391 GAGGTGGTGGCAAGAACTCTTGGCCAATGGTGATACAGGGGCCCTCTTCTTCTTCTTCT 5450
Db 5248 GACATCAATTGGCAAGAACTCTCTGGCCATGATGGCCAGGGACCTCTGTTCCTCTCTCATC 5307
Qy 5451 ACATCTACTGTGCAGCACCGAAGCCAACTCTCTGCCACAGCCAGGGGTGAGGTCTCTGCCA 5510
Db 5308 ACATCTCTCTCCAAACCCGCAACCGTCTCTCTGCCACAAATCAAAAACCAAGACTGCTGCGG 5367
Qy 5511 CTCTGGGAGAGGAGGACGAGATGTAGCCTGTGAACGGGAGCGGGTGGTCCAAAGAGCC 5570
Db 5368 CCCCTGGGGAGGAGGATGAGGATGTGGCTCAAGAGCGTGGAGCGGGTGAACCAAGGGGCT 5427
Qy 5571 ACCAGGGGAGTGTGTGGTCTGAGGAACTTTGACAAAGGTATACCTGGGGCAGAGGATG 5630
Db 5428 ACCAGGGGAGTGTGTAGTCTCTCAGGGACTTTGACCAAGGTATTACCTGGGGCAGAGGAAC 5487
Qy 5631 CCAGCTGTGACCGCTTGTGTCTGGGATTCCTCTGTGTGATGTGTGTGGGTGCTGGGT 5690
Db 5488 CCAGCTGTGATCGCTGTGTCTAGGGATCCCTCTGGGGAGTGTGTGGGCTGTGGGT 5547
Qy 5691 GTGAATGAGCAGAGGAGAGACCTCCACGTTTCGCAATGGTGAAGGGGACACATTTGGCCAGC 5750
Db 5548 GTCAACGGGGCAGGAGAGACATCACTTCCGCAATGGTGAAGGGGACACACTTGGCCAGC 5607
Qy 5751 AGGGCCAGAGGTGTGTGGCAGGCGCACAGCGTGGCCCGGGAACCCAGTGTCTGCGCACCTC 5810
Db 5608 AGTGTGAAGCAGTACTGGCAGGCGCAACAGTGGCCCAAGAGCGGTCTTCCGCAACCCGC 5667
Qy 5811 AGCATGGGATATGTCCTCAATCCGATGCCATCTTTGAGCTGTGAGCGGGCGCGAGCAC 5870
Db 5668 AGCATGGGCTACTGTCCCTCAGTCTGATGCCATCTTCCAGCTGTGACCGCGCGGGAACAT 5727
Qy 5871 CTGGAGCTGTGTGGCGGCTGCGCGGTGTCCGGAGGCGCCAGGTGGCCAGACCGCTGGC 5930
Db 5728 CTGGAACCTGTGTGTCTGCGCTGCGGGGTGCGCGAGGCCCAAGTGTGCGCACTGCGCTC 5787
Qy 5931 TCAGCCTTGGCGGCTGTGGGACTCTCATGTGTACGACAGCCGCGCTTGCAGGCACTTACAGC 5990
Db 5788 TCTGCTGTGTGTGCGCTTGGGCTTCTTAGCTATAGCAGACCAACCCGCGGGTACCTTACAGC 5847
Qy 5991 GGAGGGAAACAAAGCAAGCTGGCGAGCGCTGCGCTGTGGTGTGGGAGCCCAAGCCTGTGTG 6050
Db 5848 GGAGGCAACAAACGGAAGCTGGCGACAGCTTGTCTGTGTGGTGTGAGCCAGCTGTGGTCTC 5907
Qy 6051 TTTCTGAGCAGGCGGACCAAGGATGAGCCCAAGCGCGCGCGCTTCTTTGGNAACAGC 6110
Db 5908 TTTCTGAGCAGGCGGACCAAGGATGAGCCCAAGTGTGCGCGCGGATTTCTTTGGAAACAGC 5967
Qy 6111 TTTTGGCCGTGTGTGCGGGAGGCGGCTTCACTGTGTGTCTCACTCCCATAGCATGGAGAG 6170
Db 5968 TTGCTGTCCGTGTGTGCGGAGGCGGCTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6027
Qy 6171 TGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTGTGAATGGGCGGCTTCCGCTGTGCGG 6230
Db 6028 TGCGAACGCTCTGCAAGCGCTTGGCCATCATGTGTGAACGGGCGGCTTCCGCTGTCTGGGA 6087

QY 6231 AGCCCGCAACATCTCAAGGCGAGATTCCGCGGGTCACACATGACCCCTCGCGGTGCC 6290
DB 6088 AGCTCTCAGCATCTCAAGAGGAGTTCGCGGCTGGCCACACATGACTCTAAGGGTCCCA 6147
QY 6291 GCGCAAGGTCCTCAAGCGCGAGCGGCTTCGTCGCGGCCGAGTTCCCTCGGGTCGAGCTG 6350
DB 6148 CCGGACAGCTGAGCGCGGATAGCTTCATCAGGATCATTCTCTCGGGCTGAACTC 6207
QY 6351 CGGAGGCAATGAGGCGCGCTGCGCTTCACAGCTGCGCGCGGAGGCGCTGCGCCCTG 6410
DB 6208 CGGAGGTGACGCGAGCGCTGCGCTTCACATGCGCACGCGGGGCGAGATGACCCCTG 6267
QY 6411 GCGGCGCTTTGGAGAGCTGCGGTCGACGCGCGCAGACGCGCTGGAGGACTTTTCC 6470
DB 6268 ACAGAGTGTTCAGGGAGCTGGCTGCCAGCGGAGGCGCCACCGTGTGGAGGACTTCTCT 6327
QY 6471 GTGAGCCAGACGATGCTGGAGGAGTATTCTTGACTTCTCAAGGACAGGGGAGAGAC 6530
DB 6328 GTGAGCCAGACCTCTGGAGGAGTGTTCATATATTTCTCCAAAGACCAAGGGAGAG 6387
QY 6531 GAGGACACCGAAGACGAGAGGAGGAGTGGAGTGAGACCCCGCGCAGGCGCTGCGAG 6590
DB 6388 GAAGAGAGC--AGTCGCGAGGAGCTGAAGAGAGGAGGTTTCCAAACCTGGCGGCGAG 6444
QY 6591 CACCCGAAACGCTCAGCAGTTCCTCGATGACCTAGACACTGCCGAGACTGTCTCTGA 6650
DB 6445 CATCCGAAACGCTGACGCGGATTCCTGGAAGACCCAGCTCTGTGGAGACCATGATCTGA 6504
QY 6651 GCCTCCCTCCCTCGCGGGCGCGGGAGGCGCTTGGGAATGG 6692
DB 6505 GCATGCTGCTTGGGACTCAGTGGCAAGCTCAGACAGAGG 6546

RESULT 15
AX080493
LOCUS AX080493 6607 bp DNA linear PAT 26-FEB-2001
DEFINITION Sequence 32 from Patent WO0109314.
ACCESSION AX080493
VERSION AX080493.1 GI:13162147
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Chimini, G.
TITLE Novel uses of abca-type transporters
JOURNAL Patent: WO 0109314-A 32 08-FEB-2001.
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE; (INSERM)
(FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
LOCATION/Qualifiers
FEATURES
source 1..6607
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 57.1%; Score 3885.8; DB 6; Length 6607;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 4984; Conservative 0; Mismatches 1442; Indels 153; Gaps 8;
QY 210 ATGCGCTTTGGACACAGCTGATGCTGCTCTGGAAGAATTTTCATGTATCGCGGAGA 269
DB 1 ATGCGCTTTAGGSCACACAGCTGATGCTTCTGCTGGGAAAAATTACACCTATCGACGGAGA 60
QY 270 CAGCGGCTCAGCTCTGCTGGAATGCTGCTGGGCTCTCTCTCTCTCTCTCTCTCTCTG 329
DB 61 CAACCGATCAACTACTAGTGGAGTGTCTTTGGGCGCTCTCTCTCTCTCTCTCTCTCT 120
QY 330 GCTGTTCCGCACTCCACCGCGCCCTGGAGCACCATGAATGCCACTTCCCAAAACAGGCCA 389
DB 121 GCTGCTGCTCACTCCACCGCGCTCTGAGGATCAGAAATGCCACTTTTCCAAACAGGCCA 180

QY 390 CTGCCATCGCGGGCACCGTGCCTCGCTCCAGGGTCTCATCTGTATGTGAACAACACC 449
DB 181 TTACCATCGCGGGCACCGTGCCTCGCTCCAGGGCTTGTCTGCAACGTAACAACATCC 240
QY 450 TGTCTTTCGCGAGCTGACACGCGGCGAGGAGCCCGGCGCTGAGCAACTTTCAACGACTCC 509
DB 241 TGTCTTCCAGCACCCAAACGCTGCGGAGAGGCTGGGGTCTCTGAGTAACCTTTAAGGATTCC 300
QY 510 CTGGTCTTCCGCGTGTAGCCGATGCCGACTGTGCTGGGAGGGGCGCTAGTGCCTCAGG 569
DB 301 TTGATCTCAGGGCTCTCTGCTGATACCCGACACAGTGTCTCGGGGCGCACAGCATCCAGGAC 360
QY 570 ACCTGGCTGGCTAGGAGAGCTGATCGCCAGCTGAGGCGCTGCAACGACGACGCGCCAG 629
DB 361 ATGCTGGATGCCCTGGGAAACTGATCCCGCTGCTCAGGCGCATTTGGAGGTGGAGCACGA 420
QY 630 CTTCAACCAACCAAGCAGTCTCCACTGGAAACCAACCATGCTGATGTGCGGGAGCTGCTG 689
DB 421 CCACAGAGAGTGCACGACCGGACCACTCAAGGGTCT-----AGTGACTAAGCTTCTG 471
QY 690 ACCTCACTGCTGCGCACGGAAATCCCTGGGGTTGGCACTGGGCGCAAGCCCAAGGAGCCCTTG 749
DB 472 GAGAAAGATCCTGCAAGGGCATCCCTGGATCCTGTGCTGGTCAAGCCCAAGGATTTATG 531
QY 750 CACAGCTTGTGGAGGCGCTGAGGACCTGGGCGGAGGAGCTCTGCGGCTGCGGAGCCTG 809
DB 532 AGAAAGTTCTCAGATCTATCAGGGATCTTGGCCACGAGGCTCTGTCACATGCGCCAGCCTG 591
QY 810 GTGGAGCTTGGGAGCTGCTGACAGACCCCGAGGAGCACAGCGGCGCCCTTGGAGTTGCTG 869
DB 592 ATGGAGCTCGGAGCTTGTCTGCGAGGCGCCGAGGCTGAGCTGGTCTCTGGAGCTGGTT 651
QY 870 TCAGAGGCGCTCTGCACTGCTCAGGGAGCTTAGCAGCACAGTGGGCGCCCTTCCCTCAACTGG 929
DB 652 TCGGAGGCGCTCTGCACTGATCAAGGAGCACGACAGTCCAGGGGCGCTGTCCCTCAATTGG 711
QY 930 TACGAGGCTAGTACCTGATGAGCTGCTGGGCGGAGGAGCCAGAAATCCGCGCTTGGCAGAC 989
DB 712 TACGAGGCGCAACAGCTTAATGAGTTCTATGGGGCGGAGGTTGGCGCTTGGCGCTGAC 771
QY 990 AGCAGGCTGAGCGCGCGCTGCTCGAGGCTGATTTGGAGCGCTGACAGCCACCGCGCTGTCC 1049
DB 772 AACAGTCTCAGCGCTGCTGCTGAGTTTGTGGGACACTGGATGACCACTCTGTGTCT 831
QY 1050 CGCTCTCTGGAGAGCGCTGAAAGCTCTGATCCTCGGAGAGCTACTTTTGGACAGAT 1109
DB 832 CGGCTGCTCTGGAGGCGCTGAAAGCTTGAATCTCTCGGAAAAATTTCTTTGCACTGAC 891
QY 1110 ACACCTTTTACCGGAGGCTCATGGCGGAGGAGTGAACGGAGCTTTCGAGGAGCTCACCCCTG 1169
DB 892 ACAAACTTCACTCGGAAGCTCATGGCTCAGGTGAACAGACCTTTCGAGGAGCTGGGCTGTG 951
QY 1170 CTGAGGAGTGTCTCGGAGGCTGCTGGAGATGCTGGGACCCCGGAGCTTTCACCTTCATGAAC 1229
DB 952 TTGAGGAGCTTACACGAACTCTGCGGGGTGCTGGGAGCCCGAGATCTTCACTTCATGAAT 1011
QY 1230 GACAGTTTCCAAATGTGGCCATGCTGACGCGGCTCTCTGAGATGAGGATGAAGAGAGG 1289
DB 1012 GACAGTACCAACGCTGGCCATGCTTTCAGAGGCTTCTGATGTTGGGGGCGCACAGGCGAGG 1071
QY 1290 CAGCCAGAGCTGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1349
DB 1072 CAGCAGACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
QY 1350 AGCGGTGGCTTACAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1409
DB 1132 AGGGGTGGCTTACAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1191
QY 1410 GCGCGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1469
DB 1192 GGACAAATGATGAGTGTGTGCTCCCTGGGACCAAGCTGGAGGCTGTGCGCTCAGAGGAAGCT 1251

Db 3412 CTGGATCAGCAGCTGGAGCTCTGGGTCTCACAGGCTACGGGATCTCGGACACCAACCTG 3471
Qy 3627 GAGGAGATCTTCTGAAGGTGGAGAGTGTCTCGGACACACAGATATGGAGGTGGC 3686
Db 3472 GAGGAGATCTTCTGAAGGTGGAGAGTGTCTCGGACACACAGATATGGAGGTGGC 3525
Qy 3687 AGCTGCGGGCAGCAGCTATGACAGGCAATGCTGGCTAGACGTAACCTCGCGCTCAAG 3746
Db 3526 AGACCGCAGCTGCACCTTCGCACATGC-----ACTCCACAGCCCC 3567
Qy 3747 ATGCCGCCACAGGAGACAGAGCTGGAGAACGGGAAACAGCTGGGTACGCCCCAGAGACT 3806
Db 3568 ACAGGGCCAGAGGCATCAGTCTCGAGAAATGGGAGCTGGC----- 3608
Qy 3807 GACCAGGGCTCTGGGACAGACGCCGTGGGCGGGGTACAGGGCTGGGGCACTGACCCGCCAG 3866
Db 3609 --CCACAGGGCTTGGACACCAACGCTGCCCAAGTGCAAGGCTGGACACTTACCTGTGCAA 3666
Qy 3867 CAGCTCCAGGCCCTGTCTTCTCAAGCGCTTTCGCTTCCCGCGCGCAGCGCCGCGGCTG 3926
Db 3667 CAGCTCCGGGCTGTCTTCCAAAGCGTTTCTGCTTCCCGCAGCGCCGCGGCGCTG 3726
Qy 3927 TTGCCCCAGATCGTGTGCTGCTGCTCTTTGTGGGCTGGCCCTCGTGTTCAGGCTCATC 3986
Db 3727 TTTGCA CAGGTTGTGCTGCTGCTCTTTGTGGGCTGGCCCTGTGTTTTCAGGCTCAT 3786
Qy 3987 GTGCTCTCTTTGGGCACTACCGGCTCTGCGGCTCAGTCCCAACATGTAAGGTGCTCAG 4046
Db 3787 GTGCTCTCTTTGGGCACTACCGGCTCTGCGGCTCAGTCCCAACATGTAAGGTGCTCAG 3846
Qy 4047 GTGCTCTCTTTAGTAGGAGCGCCAGGGGACCGTGGAGTGTGCGGCTGCTCGAGGCG 4106
Db 3847 GTCTGCTTTCTAGTAGGAGTGTGCGGCTGCGGAGCCCAACCGGATGAAGCTGCTGGAGGCT 3906
Qy 4107 CTGCTGACAGGAGCAGGACTGGAGGAGCCCGCAGTGACAGATAGCTCCACAGS----- 4160
Db 3907 CTGCTAGGGAGGCTGGGCTGCAGAACCCAGTATGCAGGACAAAGATGCGAGGGATCT 3966
Qy 4161 -----TTCTCGGCAACGAAAGTTCTGCTGTAAGTGGCC 4193
Db 3967 GAGTGTACACACTCCCTAGCTTGTACTTTCAGCGTCCCTGAGGTCCCTCTGATGTGGCC 4026
Qy 4194 AAGGTCTTTGCCAGTGCACCTGACCCAGAGTCTCCATCCCGCAGCTGCCAGTGTAGC 4253
Db 4027 AGCATCTGGCCAGTGCACCTGACGCCAGATCTCCATCCCGCAGTTCGCCAATGCACT 4086
Qy 4254 CAGCCGGTGTCCGGCGCTGCTGCCGACTGCGCGCTGACGTGTGTGCTCCCGCTCCG 4313
Db 4087 CAGCTTGGAGCCCGCCCTGTTGCCAGATTGCCGCGCTGGAGCTGGGGTCCACACCC 4146
Qy 4314 CCCAGCAGTGCACCGCTCTGGGAAAGTGTTCAGAACCTGACAGGCCGGAACTGTCT 4373
Db 4147 CCCAGCCTGTGGCTGGCTTTGGGGAGTGTTCAGAACCTGACAGAACCTGAGCGGAAATGTGTCT 4206
Qy 4374 GACTTCTGTCTCAAGACTACCCGCGCTGTGGCGCAGGGCTGAAGACTTAAGAAAGTGG 4433
Db 4207 GACTTTTGTGTGAAGACATACCCAGCTGTGGCGCGAGGCTTAAGAAACAAAGAGTGG 4266
Qy 4434 GTGAATGAGGTACAGGTACGGAGGCTTCTGCTGCGGGGGCGAGACCCAGGCTGCGCCCTCG 4493
Db 4267 GTGGATGAGGTACAGATATGGGGCTTCTCCTGGAGGCGGAGATCCAGACCTGCCACA 4326
Qy 4494 GGCCAAAGTTGGGCCCTCAGTGGAGAGTGTGGGCGCTGCTGAGTCCCTCGCTCGGC 4553
Db 4327 GGGCATGAGGTGTCTCCCAATTTGGCAGAGATTTGGGCACTGCTGAGCCGCCCAACCTGGG 4386
Qy 4554 GGGCCCTCGACCTGTCTGAAACCTCACAGCTGGGCTCACAGCCTGGATGCTCAG 4613
Db 4387 AATGCGCTAGACCGTATCTTGAACACCTCAGTGTGGGCCCTTGGCCCTTGTATGCTCGG 4446
Qy 4614 GACAGTCTAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTGGCTTTGTCAAC 4673
Db 4447 AACAGCCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTGCCATGGTGGCTTTGTGAAC 4506

Qy 4674 CGAGCCAGCAACGCAATCTCTCGTCTCACTGCCCCCGAGGCCGCCCGCCACGCCAC 4733
Db 4507 CGAGCCAAACAAATGGATCTCTACATGCCCTCTTACCATCTGGTCCAGTCCGCCATGCCAC 4566
Qy 4734 AGCATCACACACACTCAACACCCCTTTGAACCTCACCAAGGAGCAGCTGTCTGAGGCTGCA 4793
Db 4567 AGCATCACTACACTCAACCATCTTTGAACCTTGACCAAGGAGCAGCTATCTGAAGCTACA 4626
Qy 4794 CTGATGGCTCTCTCGGTGAGCGTCTGCTCTCATCTGTGTGTCTTTTGTGCATGTCTTT 4853
Db 4627 CTGATAGCCTCTCTCTGTGGATGCTCTGCTCTCCATCTGTGTGTCTTTTGTGCATGTCTTT 4686
Qy 4854 GTCCCCGCGCAGCTTCACTCTTGTCTCATCTGAGGAGGAGTCAACCGAGCCAGACCACTG 4913
Db 4687 GTCCCCGCGCAGCTTCACTCTTGTCTCTCATCTGAGGAAACGATCACAGAGCAAGCATCTG 4746
Qy 4914 CAGCTCATGGGGGCGCTGTCCCCCAACCTCTACTGTGCTTTGGCAATTTTCTCTGGGACATG 4973
Db 4747 CAGCTGTGTCAGGGCTGTCCCCCAACCTCTATTTGGCTTTGGCAATTTCTCTGGGACATG 4806
Qy 4974 TGTAACTACTGTGTGTCAGCATGCTGTGTGCTCATCTTTCTGTGCTTTTCTGAGGAGG 5033
Db 4807 TGTAACTACTGTGTGTCAGCATGCTGTGTGCTCATCTTTCTGTGCTTTTCTGAGGAGG 4866
Qy 5034 GCATATGTGGCCCTGCCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5093
Db 4867 GCATATGTGGCCCTGAGAACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4926
Qy 5094 TCGATCACACCGCTCATGTATACCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5153
Db 4927 TCTATCACACCACTCATGTATACCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4986
Qy 5154 GTGGTCTCACTGTGCAATAACCTTTTATGTGGCATCAATGGAAGCATGTGCCACTTTGTG 5213
Db 4987 GTGGTCTCACTGTGCAATAACCTTTTATGTGGCATCAATGGAAGCATGTGCCACTTTGTG 5046
Qy 5214 CTTGAGCTCTTCTCTCATCAGAGCTGACAGGAGTGCAGCGGATCTTGAACAGGTCTTTC 5273
Db 5047 CTAGAACTGCTTTTTCAGATCAGAACTTGCAGAAAGTGCAGCGGATCTTGAACAGGTCTT 5106
Qy 5274 CTTATCTTCCCCCACTTCTGTGTGGCCCGGGGCTCATTTGACATGTGTGGGAAACAGGCC 5333
Db 5107 CTTATTTTCCCCCACTTGTGCTTGGCCGAGGGCTCATTTGACATGTGTGGGAAACAGGCC 5166
Qy 5334 ATGGCTGATGCTTTGAGCGCTTTGGAGACAGGCACTTCACTCACTCTGCTGCTGGAG 5393
Db 5167 ATGGCAGATGCTTTGAGCGCTTTAGGAGACAAAGCAATTTTCACTCACTCTGCTGGAC 5226
Qy 5394 GTGGTGGCAAGAACCTTGTGGCCATGTGTGATACAGGGGCCCTCTTCTCTCTCTTTCACA 5453
Db 5227 ATCAITTTGGCAAGAACCTTGTGGCCATGTGTGATGAGGAGCTCTGTCTCTCTCTTTCACA 5286
Qy 5454 CTACTCTGTCAGCACCAAGCCAACTCTCTGCAACAGGCCAGGCTGAGGTCTCTGCACTC 5513
Db 5287 CTCCTGCTCCACACCGCAACCTCTCTGCAACATCAAAACCAAGACTGCTGCGGCC 5346
Qy 5514 CTGGGAGAGGAGCAAGGATGTAGCCCTGTGAACCGGAGCGGGTGTTCAGAGAGGCCACC 5573
Db 5347 CTGGGAGAGGAGATGAGGATGTGGCTCAAGAGCGTGTGAGCGGGTGCACCAAGGGGGTATCC 5406
Qy 5574 CAGGGGATGTGTGTGTGTGAGGAATTTGACCAAGGTATACCGTGGGAGAGGATGCCA 5633
Db 5407 CAGGGGATGTGTGTGTGTGTGAGGAATTTGACCAAGGTATACCGTGGGAGAGGATGCCA 5466
Qy 5634 GCTGTGACCGCTGTGTGCTGGGATTTCCCTCTGTGTGAGTGTGTGTGTGTGTGTGTGTG 5693
Db 5467 GCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5526
Qy 5694 AATGAGCAGGAGGAGAGCGTCCAGCTTTTGTGATGAGTGTGTGTGTGTGTGTGTGTGTG 5753
Db 5527 AACGGGCGAGGAGAGAGCATCCACTTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5586

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:41:30 ; Search time 2868 Seconds
(without alignments)
12453.647 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctccagggcgccgcctccc.....cgtgctggtgaaaaa 6804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1:	Geneseq1980s:*
2:	Geneseq1990s:*
3:	Geneseq2000s:*
4:	Geneseq2001as:*
5:	Geneseq2001bs:*
6:	Geneseq2002as:*
7:	Geneseq2002bs:*
8:	Geneseq2003as:*
9:	Geneseq2003bs:*
10:	Geneseq2003cs:*
11:	Geneseq2003ds:*
12:	Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6804	100.0	6804	8	ABX14666 Human CDN
2	6762.4	99.4	6791	6	AAL44693 Human tra
3	6717.8	98.7	6768	6	AA519207 DNA encod
4	6717.8	98.7	6768	10	Add37428 Human tra
5	6710	98.6	6795	6	ABL57810 Human ABC
6	6478.4	95.2	6522	4	ABX08706 Human PD-
7	6398.4	94.0	6432	10	Add37430 Human tra
8	6036.8	88.7	6327	8	ABX72257 Human NOV
9	5933.2	87.2	6696	8	ABX95283 CDNA enco
10	5833.2	85.7	6651	8	ABX95284 CDNA enco
11	5529	81.3	5811	4	AA05626 Human sec
12	5529	81.3	5811	8	ADA40555 Human sec
13	5529	81.3	5811	10	ADA56706 Gene enco
14	5458.2	80.2	5669	4	AA508707 Human PD-
15	4843.4	71.2	5859	8	ABX95285 Human ABC
16	4381.6	64.4	4653	8	ABX14667 Human CDN
17	3902.4	57.4	6633	8	ABX14665 Mouse CDN
18	3895.8	57.1	6607	4	AAF54812 Nucleotid
19	3447.4	50.7	3635	4	AAF56389 ABC trans
20	3303	48.5	5762	4	AAF54792 Nucleotid
21	3129.8	46.0	3437	5	AA573965 DNA encod

22	3039.8	44.7	4413	5	AA583715	AA583715 DNA encod
23	2861.8	42.1	2911	3	AAZ94751	AAZ94751 Human ATP
24	1808.6	26.6	1879	11	ADM03146	ADM03146 Human cDN
25	1740.6	25.6	6786	9	ACC84923	ACC84923 Human ABC
26	1739	25.6	10412	9	ACC84922	ACC84922 Human ABC
27	1739	25.6	10412	11	ADP65742	ADP65742 Human ATP
28	1739	25.6	10412	11	ADP64975	ADP64975 Human ATP
29	1739	25.6	10412	12	ADM41251	ADM41251 Human ATP
30	1737.4	25.5	7860	4	AAF92835	AAF92835 Human ABC
31	1737.4	25.5	7860	5	AAF83826	AAF83826 Human ABC
32	1737.4	25.5	10365	10	ADD94068	ADD94068 Human ATP
33	1737.4	25.5	10381	10	ADD94069	ADD94069 Human ATP
34	1737.4	25.5	10423	10	ADD94070	ADD94070 Human ATP
35	1737.4	25.5	10442	4	AAF24680	AAF24680 Nucleotid
36	1737.4	25.5	10442	4	AAF24702	AAF24702 Nucleotid
37	1735.8	25.5	9741	6	ABL58146	ABL58146 Human ABC
38	1735.8	25.5	9870	6	ABN99319	ABN99319 Polymorph
39	1735.8	25.5	9870	6	ABN99316	ABN99316 Polymorph
40	1735.8	25.5	9870	6	ABN99315	ABN99315 Polymorph
41	1735.8	25.5	9870	6	ABN99308	ABN99308 Polymorph
42	1735.8	25.5	9870	6	ABN99318	ABN99318 Polymorph
43	1735.8	25.5	10474	4	AAF24685	AAF24685 Nucleotid
44	1735.8	25.5	10474	4	AAF24686	AAF24686 Nucleotid
45	1735.8	25.5	10474	4	AAF24708	AAF24708 Nucleotid

ALIGNMENTS

RESULT 1						
ABX14666	standard; CDNA; 6804 BP.					
AC	ABX14666;					
DT	10-MAR-2003 (first entry)					
DE	Human cDNA encoding ATP-binding cassette transporter-like protein.					
XX	Human; ATP-binding cassette transporter-like protein; ABCCL;					
KW	lipid transport; cardiovascular disease; hypertriglyceridaemia;					
KW	atherosclerosis; hypercholesterolaemia; Tangier disease; dyslipidaemia;					
KW	nervous system disorder; Stargardt disease; degenerative disorder;					
KW	inflammatory retinopathy; cystic fibrosis; multidrug resistance;					
KW	lymphoid condition; myeloid cell condition; AIDS; lymphoma;					
KW	acquired immunodeficiency disorder; leukaemia; neutropaenia; anaemia;					
KW	autoimmune disease; thyroid disorder; hyperthyroidism; hypothyroidism;					
KW	hypothalamus disorder; obesity; diabetes; reproductive disorder;					
KW	energy balance disorder; peripheral neuropathy; myelinopathy; ss; gene;					
KW	axonopathy; autoimmune disease; inflammatory disease; multiple sclerosis					
XX	Homo sapiens.					
XX	Key					
XX	Location/Qualifiers					
FT	CDS	210..6650				
FT		/*tag= a				
FT		/product= "ABCL"				
FT	sig_peptide	210..347				
FT		/*tag= b				
FT	mat_peptide	348..6647				
FT		/*tag= c				
FT		/label= Mature_ABCL				
XX	US2002127647-A1.					
PN	XX					
XX	XX					
PD	12-SEP-2002.					
XX	XX					
PF	28-NOV-2001; 2001US-00995542.					
XX	XX					
PR	28-NOV-2000; 2000US-0253520P.					
XX	XX					
PA	(SHUT/) SHUTTER J.					
PA	(ULIA/) ULIAS L.					

Db 1441 AGCTGAGGGGACCCCTCAGAGGACGCCCTGGTGTGGGGCCCTGCAACTGTGCTGGCGG 1500
Qy 1501 AACATGATTTCTGGGCGGCGTGTCTTCTTGGGACCTGAGAGCTCTTTCAGACCCACAG 1560
Db 1501 AACATGATTTCTGGGCGGCGTGTCTTCTTGGGACCTGAGAGCTCTTTCAGACCCACAG 1560
Qy 1561 AGCACCCACCCAGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
Db 1561 AGCACCCACCCAGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
Qy 1621 AGCTGGTCAAGAGGACCAATAGATCAGGACAGGTTTTTGGGACCCCTGGCCAGCGCGG 1680
Db 1621 AGCTGGTCAAGAGGACCAATAGATCAGGACAGGTTTTTGGGACCCCTGGCCAGCGCGG 1680
Qy 1681 ACCCCCTGACCGACCTGCGCTACGTGTGGGCGGCTTCTGTGTACTGTGCAAGACCTGGTGG 1740
Db 1681 ACCCCCTGACCGACCTGCGCTACGTGTGGGCGGCTTCTGTGTACTGTGCAAGACCTGGTGG 1740
Qy 1741 AGGTGACAGCGCTCGCGTGTCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1800
Db 1741 AGCGTGCAGCGCTCGCGTGTCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1800
Qy 1801 AGATGCCCTATCCGCTGTATGTGGACGAGCGTGTCTCGGTGTGTGAGCGGCTCGCTGC 1860
Db 1801 AGATGCCCTATCCGCTGTATGTGGACGAGCGTGTCTCGGTGTGTGAGCGGCTCGCTGC 1860
Qy 1861 CGCTCTTCTGACCGCTGGCTGGATCTACTCCGTGACACTGACAGTGAAGCGGCTGGTGC 1920
Db 1861 CGCTCTTCTGACCGCTGGCTGGATCTACTCCGTGACACTGACAGTGAAGCGGCTGGTGC 1920
Qy 1921 GGGAGAGAGAGCGCGGTGCGGGACACATGCGCGCCATGCGGCTCAGCGCGCGGCTGC 1980
Db 1921 GGGAGAGAGAGCGCGGTGCGGGACACATGCGCGCCATGCGGCTCAGCGCGCGGCTGC 1980
Qy 1981 TCTGCTAGGCTGTCTCAGCTGCTGCGGCGGCTTCTGCTGAGCGCGGCTGCTGG 2040
Db 1981 TCTGCTAGGCTGTCTCAGCTGCTGCGGCGGCTTCTGCTGAGCGCGGCTGCTGG 2040
Qy 2041 TTCTGGTGTCTCAAGCTGGGGACATCTCTCCCTACAGCACCCCGGGCGTGTCTTCTGT 2100
Db 2041 TTCTGGTGTCTCAAGCTGGGGACATCTCTCCCTACAGCACCCCGGGCGTGTCTTCTGT 2100
Qy 2101 TCTTGGAGCCTTTCGCGTGGGCAACCGGTGAACCGAGCTTCTGCTCAGCGGCTTCTTCT 2160
Db 2101 TCTTGGAGCCTTTCGCGTGGGCAACCGGTGAACCGAGCTTCTGCTCAGCGGCTTCTTCT 2160
Qy 2161 CCGCGGCAACCTGGCTGGGCTGCGGCGGCTGCGGCTTCTTCTCCTCTACCTGGCCT 2220
Db 2161 CCGCGGCAACCTGGCTGGGCTGCGGCGGCTGCGGCTTCTTCTCCTCTACCTGGCCT 2220
Qy 2221 ACGTGTGTGTGGGCGGACCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280
Db 2221 ACGTGTGTGTGGGCGGACCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280
Qy 2281 TGCTGTGCGGCTTTCGCTTTCGCTGCGAGAGCTTGGCTGCTGCTGGAGGAGCAGG 2340
Db 2281 TGCTGTGCGGCTTTCGCTTTCGCTGCGAGAGCTTGGCTGCTGCTGGAGGAGCAGG 2340
Qy 2341 GCGAGGCGGCGAGTGGGACAAAGTGGGACCCCGGCTTACGGCAGAGCTTTCAGCCTGG 2400
Db 2341 GCGAGGCGGCGAGTGGGACAAAGTGGGACCCCGGCTTACGGCAGAGCTTTCAGCCTGG 2400
Qy 2401 CCCAGTCTTGGGCTTCTGCTGTGGACGCGGCGCTCTACGGGCTGCGCACCTGGTACC 2460
Db 2401 CCCAGTCTTGGGCTTCTGCTGTGGACGCGGCGCTCTACGGGCTGCGCACCTGGTACC 2460
Qy 2461 TGGAGCTGTGTGCGGCGGAGTACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Db 2461 TGGAGCTGTGTGCGGCGGAGTACGGGATCCCTGAAACCATGGAATTTTCTTTTCGGA 2520
Qy 2521 GGAGTACTGTGCGGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580
Db 2521 GGAGTACTGTGCGGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580

Qy 2581 ACCCAAAGTGTGTTAGAGAGGCAACGCGCGGCGCTGAGTCTTGGCGTATCCGTTGCA 2640
Db 2581 ACCCAAAGTGTGTTAGAGAGGCAACGCGCGGCGCTGAGTCTTGGCGTATCCGTTGCA 2640
Qy 2641 GCCTGGAGAGGCTTTCTGGAAGCCCGAGCGGCGGCTGCGGGGCTCAGCCTGGACT 2700
Db 2641 GCCTGGAGAGGCTTTCTGGAAGCCCGAGCGGCGGCTGCGGGGCTCAGCCTGGACT 2700
Qy 2701 TCTACAGGGGCACATCACCGCTTCTGGGCGCAACGCGGCGGCGGCAAGACACACCC 2760
Db 2701 TCTACAGGGGCACATCACCGCTTCTGGGCGCAACGCGGCGGCGGCAAGACACACCC 2760
Qy 2761 TGTCCATCTTGTAGTGGGCTTCTCCACCCAGTGGTCTGCTGCTTCTTCTGGGCGGCAAG 2820
Db 2761 TGTCCATCTTGTAGTGGGCTTCTCCACCCAGTGGTGGTCTGCTGCTTCTTCTGGGCGGCAAG 2820
Qy 2821 AGCTCGCTCAGACATGCGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTTCTCAGTACAAG 2880
Db 2821 AGCTCGCTCAGACATGCGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTTCTCAGTACAAG 2880
Qy 2881 TGCTGTTTGAATGCTGACCGTGGACGAGCAGCTGCTGCTTCTTCTGGGCGGCTTCTTCTG 2940
Db 2881 TGCTGTTTGAATGCTGACCGTGGACGAGCAGCTGCTGCTTCTTCTGGGCGGCTTCTTCTG 2940
Qy 2941 TGAGTGGCGCTGTAGTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTTCTGGGCGGCT 3000
Db 2941 TGAGTGGCGCTGTAGTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTTCTGGGCGGCT 3000
Qy 3001 CCAAAGAGTGTGAGACTCGGCACTCTCTGCTGGGATGCAACGGAAGCTGTCCGTTGG 3060
Db 3001 CCAAAGAGTGTGAGACTCGGCACTCTCTGCTGGGATGCAACGGAAGCTGTCCGTTGG 3060
Qy 3061 CCATGCTTGTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGG 3120
Db 3061 CCATGCTTGTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGG 3120
Qy 3121 ATCTGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGG 3180
Db 3121 ATCTGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGGCTTCTGGGCGG 3180
Qy 3181 TGATCTCTCCACCACTGATGAGGAGAGCTGCTGGGAGACCGTGGGCTGTGG 3240
Db 3181 TGATCTCTCTCCACCACTGATGAGGAGAGCTGCTGGGAGACCGTGGGCTGTGG 3240
Qy 3241 TGGCAGGTTGGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
Db 3241 TGGCAGGTTGGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
Qy 3301 CCGGCTACTACTGAGCTGGTGAAGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGG 3360
Db 3301 CCGGCTACTACTGAGCTGGTGAAGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGG 3360
Qy 3361 ACATGACATGAGGGGAGTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420
Db 3361 ACATGACATGAGGGGAGTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420
Qy 3421 GCAGAGTGGGACCTTCTGAGCTGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 3480
Db 3421 GCAGAGTGGGACCTTCTGAGCTGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 3480
Qy 3481 TGGTGGAGGAGCTGCCACAGAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3540
Db 3481 TGGTGGAGGAGCTGCCACAGAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3540
Qy 3541 GCAGCTTTCGCGACACTCTTTCGAGAGCTTAGACACGCGGCTGGCGGAGCTGAGGCTCACTG 3600
Db 3541 GCAGCTTTCGCGACACTCTTTCGAGAGCTTAGACACGCGGCTGGCGGAGCTGAGGCTCACTG 3600
Qy 3601 GCTACGGGATCTCCGACACAGGCTTCGAGGAGTCTTCTTCTGAGGAGTGGTGGAGGAGTGG 3660
Db 3601 GCTACGGGATCTCCGACACAGGCTTCGAGGAGTCTTCTTCTGAGGAGTGGTGGAGGAGTGG 3660

QY	3661	CTGCGGACAGATATGAGGATGGCAGCTGCGGGCAGCACCTATGACAGGCAATTCGTCG	3720
Db	3661	CTGCGGACAGATATGAGGATGGCAGCTGCGGGCAGCACCTATGACAGGCAATTCGTCG	3720
QY	3721	GCCTAGACGTAACCTTCGGCTCAAGATGCGCGCACAGGAGACAGCGCTGGAGAACGGGG	3780
Db	3721	GCCTAGACGTAACCTTCGGCTCAAGATGCGCGCACAGGAGACAGCGCTGGAGAACGGGG	3780
QY	3781	AACAGCTGGGTCHAGCCCCAGAGACTGACAGGGCTCTGGGCGACAGCCGCTGGGCGGGG	3840
Db	3781	AACAGCTGGGTCHAGCCCCAGAGACTGACAGGGCTCTGGGCGACAGCCGCTGGGCGGGG	3840
QY	3841	TACAGGCTGGGCACTGACCGCGCAGCAGCTCCAGGCGCTGCTTCTCAAGCGCTTTCGTCG	3900
Db	3841	TACAGGCTGGGCACTGACCGCGCAGCAGCTCCAGGCGCTGCTTCTCAAGCGCTTTCGTCG	3900
QY	3901	TTGCCCGCGCAGCGCGCGGGCGCTGTTGCGGCCAGATCGTCTGCGCTGCCCTCTTTGTGG	3960
Db	3901	TTGCCCGCGCAGCGCGCGGGCGCTGTTGCGGCCAGATCGTCTGCGCTGCCCTCTTTGTGG	3960
QY	3961	GCTGGGCGCTGCTTTCAGCGCTCATCGTGCCTCTCTTTGCGGCACTACCGGCTCTGCGGC	4020
Db	3961	GCTGGGCGCTGCTTTCAGCGCTCATCGTGCCTCTCTTTGCGGCACTACCGGCTCTGCGGC	4020
QY	4021	TCAGTCCACCATGTACGCTGCTCAGGTGTCTTCTTTCAGTGAGGAGCGCCACAGGGGACC	4080
Db	4021	TCAGTCCACCATGTACGCTGCTCAGGTGTCTTCTTTCAGTGAGGAGCGCCACAGGGGACC	4080
QY	4081	CTGAGCTGCGCGGCTGCTCGAGGCGCTGCTGACGAGGAGGAGCTGAGGAGCGCCCGAC	4140
Db	4081	CTGAGCTGCGCGGCTGCTCGAGGCGCTGCTGACGAGGAGGAGCTGAGGAGCGCCCGAC	4140
QY	4141	TGCAGATAGTCTCCACAGGTTCTCGGCACAGAACTTCTGCTGAAGTGCGCCAGGTCT	4200
Db	4141	TGCAGATAGTCTCCACAGGTTCTCGGCACAGAACTTCTGCTGAAGTGCGCCAGGTCT	4200
QY	4201	TGGCCAGTGGCAACTGGACCCCGAGTCTCCATCCCGCAGCTGCGAGTGTAGCGAGCCCG	4260
Db	4201	TGGCCAGTGGCAACTGGACCCCGAGTCTCCATCCCGCAGCTGCGAGTGTAGCGAGCCCG	4260
QY	4261	GTCGCCCGCGCTGCTGCGGACTGCGCGGTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4320
Db	4261	GTCGCCCGCGCTGCTGCGGACTGCGCGGTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4320
QY	4321	CAGTGACCGGCTCTGGGGAGTGTTCAGAACCTGACAGGCGCTGAAAGTGGGTGAATG	4380
Db	4321	CAGTGACCGGCTCTGGGGAGTGTTCAGAACCTGACAGGCGCTGAAAGTGGGTGAATG	4380
QY	4381	TGGTCAAGACTACCGCGCTGCTGCGGCGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG	4440
Db	4381	TGGTCAAGACTACCGCGCTGCTGCGGCGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG	4440
QY	4441	AGGTCAAGTACGAGGCTTCTCGCTGGGGGCGGAGACCCAGGCGCTGCGGCGCGGCGG	4500
Db	4441	AGGTCAAGTACGAGGCTTCTCGCTGGGGGCGGAGACCCAGGCGCTGCGGCGCGGCGG	4500
QY	4501	AGTTGGGCGGCTCAGTGAGAGTGTGGGCGCTGCTGAGTCCCTGCGGCGGCGGCGG	4560
Db	4501	AGTTGGGCGGCTCAGTGAGAGTGTGGGCGCTGCTGAGTCCCTGCGGCGGCGGCGG	4560
QY	4561	TCGACCGTGTCTCAAAAACCTACAGCTGGGCTGCTGAGTCCCTGCGGCGGCGGCGG	4620
Db	4561	TCGACCGTGTCTCAAAAACCTACAGCTGGGCTGCTGAGTCCCTGCGGCGGCGGCGG	4620
QY	4621	TCAGATCTGTTTCAACAAAGGCTGGCACTCCATGTTGGGCTTTGTCAACCGAGCCA	4680
Db	4621	TCAGATCTGTTTCAACAAAGGCTGGCACTCCATGTTGGGCTTTGTCAACCGAGCCA	4680
QY	4681	GCAACGAAATCTCCGCTGCTCACTGCGCCCGGCGGCGGCGGCGGCGGCGGCGG	4740
Db	4681	GCAACGAAATCTCCGCTGCTCACTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGG	4740
QY	4741	CCACACTCAACCCCTTGAACCTACCAAGGAGGAGCTGCTGAGGCTGCACTGATGG	4800

Db	4741	CCACACTCAACCCCTTGAACCTCACCAAGGAGCAGCTGTCTGAGGCTGCACTGATGG	4800
QY	4801	CCTCTCGTGGTGGAGCTGCTCGTCTCCATCTGTGTGGTCTTTTGGCATGTCTTTTGTCCGG	4860
Db	4801	CCTCTCGTGGTGGAGCTGCTCGTCTCCATCTGTGTGGTCTTTTGGCATGTCTTTTGTCCGG	4860
QY	4861	CCAGCTTCACTCTGTCTCTCATTTGAGGAGGAGCTCACCCGAGCAAGCACTCCAGCTCA	4920
Db	4861	CCAGCTTCACTCTGTCTCTCATTTGAGGAGGAGCTCACCCGAGCAAGCACTCCAGCTCA	4920
QY	4921	TGGGGGCGCTGTCCCGCACCTCTACTTGGCTTGGCAACTTTCTCTGGGACATGTGTAAC	4980
Db	4921	TGGGGGCGCTGTCCCGCACCTCTACTTGGCTTGGCAACTTTCTCTGGGACATGTGTAAC	4980
QY	4981	ACTTGTGTCAGCATGATGCTGCTGCTCATTTCTTCTGCGCTTCCAGCAGAGGCGCATG	5040
Db	4981	ACTTGTGTCAGCATGATGCTGCTGCTCATTTCTTCTGCGCTTCCAGCAGAGGCGCATG	5040
QY	5041	TGGCCCTTGGCAACCTTCTCTCTGCTGCTGCTTCTACTGATGCTGCTGCTGCTGCTG	5100
Db	5041	TGGCCCTTGGCAACCTTCTCTCTGCTGCTGCTTCTACTGATGCTGCTGCTGCTGCTG	5100
QY	5101	CACCGCTCATGTACCCAGCGCTCTTCTTCTTCTCGTGGCCAGCAGACCTATGTGGTGC	5160
Db	5101	CACCGCTCATGTACCCAGCGCTCTTCTTCTTCTTCTCGTGGCCAGCAGACCTATGTGGTGC	5160
QY	5161	TCACTGTCATAAACCTCTTTATTTGGCATCAATGGAAGCATGCGCACTTTTGTGCTGAGC	5220
Db	5161	TCACTGTCATAAACCTCTTTATTTGGCATCAATGGAAGCATGCGCACTTTTGTGCTGAGC	5220
QY	5221	TCTTCTCTGATCAGAGCTGACGAGGTCAGCGGATCTTTGAACAGGCTCTTCTTATCT	5280
Db	5221	TCTTCTCTGATCAGAGCTGACGAGGTCAGCGGATCTTTGAACAGGCTCTTCTTATCT	5280
QY	5281	TCCCCCACTTCTCTTGGGCGGGGCTCATTTGACATGCTGTCGGAACACAGGCTG	5340
Db	5281	TCCCCCACTTCTCTTGGGCGGGGCTCATTTGACATGCTGTCGGAACACAGGCTG	5340
QY	5341	ATGCTTTGAGCGCTTGGGAGACAGGAGTTCAGTCAACCCCTGCGTGGGAGGTGCTG	5400
Db	5341	ATGCTTTGAGCGCTTGGGAGACAGGAGTTCAGTCAACCCCTGCGTGGGAGGTGCTG	5400
QY	5401	GCAAGAACCTTCTGGCCATGTCATACAGGGGCGCTTCTTCTTCTTCTTCTTCTTCTTCT	5460
Db	5401	GCAAGAACCTTCTGGCCATGTCATACAGGGGCGCTTCTTCTTCTTCTTCTTCTTCTTCT	5460
QY	5461	TGCAGCACCGAAGCCAACTCTCTGCCACAGCCCGAGGTGAGGTCTCTGCGCACTCTG	5520
Db	5461	TGCAGCACCGAAGCCAACTCTCTGCCACAGCCCGAGGTGAGGTCTCTGCGCACTCTG	5520
QY	5521	AGGAGGACGAGGATGTAGCCCTGGAACGGGAGCGGTGCTCAAGAGACCCAGCGGGG	5580
Db	5521	AGGAGGACGAGGATGTAGCCCTGGAACGGGAGCGGTGCTCAAGAGACCCAGCGGGG	5580
QY	5581	ATCTGTTGTCGTCAGGACCTTCACAAAGTATACCGTGGGACAGGATGCCAGCTG	5640
Db	5581	ATCTGTTGTCGTCAGGACCTTCACAAAGTATACCGTGGGACAGGATGCCAGCTG	5640
QY	5641	ACCGCTTGTCCCTGGGGATTTCCCGCTGCTGAGTGTGTTTGGGCTGCTGAGTGGAG	5700
Db	5641	ACCGCTTGTCCCTGGGGATTTCCCGCTGCTGAGTGTGTTTGGGCTGCTGAGTGGAG	5700
QY	5701	CAGGGAAGAGCTCCATTTTCGATGAGTGGTGGGAGACATTTGGCCAGAGGGCGAGG	5760
Db	5701	CAGGGAAGAGCTCCATTTTCGATGAGTGGTGGGAGACATTTGGCCAGAGGGCGAGG	5760
QY	5761	CTGTGTCGAGGACACAGCGTGGCGGGAAACCCAGTGTCTGCGCACTCAGCATGGAT	5820
Db	5761	CTGTGTCGAGGACACAGCGTGGCGGGAAACCCAGTGTCTGCGCACTCAGCATGGAT	5820
QY	5821	ACTGCCCTCAATCCGATGCCATCTTTGAGTGTCTGAGCGGCGCGGAGCACTGAGGCTG	5880

Db	5821	ACTGCCCTCAATCCGATGCCAATCTTTGAGCTGCTGA	CGGGCCGGAGGACACTCTGGAGCTGC	5888
Qy	5881	TTGCGCGCCTCGCGCGGTGTCCCGGAGGCCACAGGT	TGCCAGACCGCTGGCTCAGGCCCTCG	5940
Db	5881	TTGCGCGCCTCGCGCGGTGTCCCGAGGCCACAGGT	TGCCAGACCGCTGGCTCAGGCCCTCG	5940
Qy	5941	CGCGTCTGGGACTCTCATGTGTACCCAGACCGGCC	CTGACGGACCTTACAGCGGAGGGAACA	6000
Db	5941	CGCGTCTGGGACTCTCATGTGTACCCAGACCGGCC	CTGACGGACCTTACAGCGGAGGGAACA	6000
Qy	6001	AACGCAAGCTGGCGAGCGGCCCTGTGGCTGTGGT	TGGGACCCAGCCGTCGTCTTCTGACG	6060
Db	6001	AACGCAAGCTGGCGAGCGGCCCTGTGGCTGTGGT	TGGGACCCAGCCGTCGTCTTCTGACG	6060
Qy	6061	AGCCGACACACAGGCATGGAACCCACAGCGCGCG	CGCTTCTTGGAAACAGCCTTTTGGCCG	6120
Db	6061	AGCCGACACACAGGCATGGAACCCACAGCGCGCG	CGCTTCTTGGAAACAGCCTTTTGGCCG	6120
Qy	6121	TGGTGCGGGAGGGCCGTTCAGTGATGCTCACTCC	ATAGCATGGAGGAGTGTGAGCGC	6180
Db	6121	TGGTGCGGGAGGGCCGTTCAGTGATGCTCACTCC	ATAGCATGGAGGAGTGTGAGCGC	6180
Qy	6181	TCTGCTCGCGCTAGCCATCATGGTGAATGGCGGT	TCGCTGGCGAGCCCGCAAC	6240
Db	6181	TCTGCTCGCGCTAGCCATCATGGTGAATGGCGGT	TCGCTGGCGAGCCCGCGCAAC	6240
Qy	6241	ATCTCAAGGCGAGATTCGCGCGGGTCAACA	CTGACCTTGGCGCGCAAGGT	6300
Db	6241	ATCTCAAGGCGAGATTCGCGCGGGTCAACA	CTGACCTTGGCGCGCAAGGT	6300
Qy	6301	CCAGCCGGCAGCGGCCCTTGTGGCGGCCAGATT	CCCTGGGTGGAGCTCGCGAGGCAC	6360
Db	6301	CCAGCCGGCAGCGGCCCTTGTGGCGGCCAGATT	CCCTGGGTGGAGCTCGCGAGGCAC	6360
Qy	6361	ATGGAGCGCGCTCGCTTCCAGCTGCGCGGGAGG	CGCTGGCCCTGGCGCGCTCT	6420
Db	6361	ATGGAGCGCGCTCGCTTCCAGCTGCGCGGGAGG	CGCTGGCCCTGGCGCGCTCT	6420
Qy	6421	TTGGAGAGCTGGCGGTGCACCGCGCAGAGCA	CGCGCTGGAGGACTTTTCCGTGAGCCAGA	6480
Db	6421	TTGGAGAGCTGGCGGTGCACCGCGCAGAGCA	CGCGCTGGAGGACTTTTCCGTGAGCCAGA	6480
Qy	6481	CGATGCTGGAGGAGTATTCTTGTA	CTTCTTCAAGGACGAGGGGAGGACGAGGACCG	6540
Db	6481	CGATGCTGGAGGAGTATTCTTGTA	CTTCTTCAAGGACGAGGGGAGGACGAGGACCG	6540
Qy	6541	AAGAGCAAGAGGAGGACGAGTGGAGTGGAC	CCCGCGCCAGGCGCTCGACGACCCCAAAC	6600
Db	6541	AAGAGCAAGAGGAGGACGAGTGGAGTGGAC	CCCGCGCCAGGCGCTCGACGACCCCAAAC	6600
Qy	6601	CGCTCAGCCAGTTCCTCGATGACCCATAGCA	CTGCGAGACTGTGCTCTGAGCCCTCCCTCC	6660
Db	6601	CGCTCAGCCAGTTCCTCGATGACCCATAGCA	CTGCGAGACTGTGCTCTGAGCCCTCCCTCC	6660
Qy	6661	CCTCGGGGCGCGGGAGGCCCTTGGGAATGGCA	AGGGCAAGCTAGAGTGCCTTAGGAGCC	6720
Db	6661	CCTCGGGGCGCGGGAGGCCCTTGGGAATGGCA	AGGGCAAGCTAGAGTGCCTTAGGAGCC	6720
Qy	6721	CTGGACTCAGGCTGGCAGAGGGGCTGTGCC	CTTGGAGAAAATAAGAGAGGCTTGAGAG	6780
Db	6721	CTGGACTCAGGCTGGCAGAGGGGCTGTGCC	CTTGGAGAAAATAAGAGAGGCTTGAGAG	6780
Qy	6781	AAGCCGTGCTGGTGA	AAAAA 6804	
Db	6781	AAGCCGTGCTGGTGA	AAAAA 6804	

RESISTANCE

RESULI 2
AAI.44693

AAI44693
ID AAI44693 standard: cDNA: 6791 BP.

XY
IDAA
AC
AAL44693:

XX XX

03-MAY-2002 (first entry)

Human transporter and ion channel TRICH-27 cDNA.

Human; transporter and ion channel; TRICH; transport disorder; neurological disorder; muscle disorder; immunological disorder; cell proliferative disorder; neuroprotective; nootropic; cerebroprotective; immunosuppressive; cyostatic; respiratory; gene therapy; gene; ss.

Homo sapiens.

WO200204520-A2

17-JAN-2002

05-ПП.-2001: 2001W0-IIS021448

07-:III.-2000: 2000IS-0216547D

07=JUL=2000; 2000US=0216347P;
14=JUL=2000; 2000US=0218232P

14-JUL-2000; 2000US-0218Z3ZF.
21-JUL-2000: 2000US-022011ZP.

28-JUL-2000; 2000US-0221839P;

(INCY-) INCYTE GENOMICS INC.

Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L; Burford N, Greene BD, Sanjanwala MS, Baughn MR, Yao MG, Yang J; Paterson C, Gandhi AR, Hafalia AWA, Triboulet CM, Wallia NK; Au-Young J, Walsh RT, Rankumar J, Lu Y, Lu DM, Azimzal Y, Lal P; Elliott VS, Nguyen DB, Xu Y, Seilhamer JJ, Borowsky ML, Khan PA; Kearney L, Thangavelu K, Das D, Pollick JL;

WPI: 2002-205969/26.

New human transporters and ion channel polypeptides for diagnosing, treating or preventing transport, neurological, muscle, immunological and cell proliferative disorders.

Claim 5: Page 224-225: 230pp: English:

The present invention provides the protein and coding sequences of a number of human transporter and ion channel proteins, designated TRICH-1-TRICH-32. The sequences can be used in the treatment of transport, neurological, muscle, immunological and cell proliferative disorders. The present sequence is a coding sequence of the invention

Sequence 6791 BP: 1137 A: 2202 C: 2184 G: 1268 T: 0 U: 0 Other:

99.4%: Score 6762.4: DB 6: Length 6791:

Best Local Similarity 99.8%; Pred. No. 0;

Matches 6780; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 12 CGCGTCCCTGCCTGCTGGCGGAGGGAAGCGGCAAGAGCTGCGGAGCCCTGGAA 71

7

D_b 1 CGCGTCCCTGCTGCTGGCGGAGGAAGCGCAAGAGCTGCGGAGCCCCCTGGAA 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 72 GAGCTTCAGGAACCCCTGCGCTGTGGATAAAGGAATGAGGTT CAGAAAGGGGCAGGGAG 131

61 CAGCTTCCACGACCCCTGCGCTGCGCTATACGACATGACGCTTACGACACCGC 119

DD GT GAGCTTCCAGGGAACCTTGCCTGTGGGATTAAGGAAATGAGGTTCAGATTAAGGGCA-TGAG 113

132 TTGCCCGCAGCCGCACCGCAGTCTTCAGCCCGACCGTTGTCTCTGACCTCTCTGCCCCGT 191

Db 120 TTGCCCCGAGCCGACCGACGTCTTCAGCCCCGACCGTTGTCTCTGTCCCCGT 179

Db 2460 TGCCACGCGCAGTACGGGATCCTGAAACCATGGAAATTTTCTCTTTTCGGAGGAGCTACTGG 2519
Qy 2532 TCGGACCTTCGCCCCCCCCAAGAGTCCAGCCCTTCGCCCCACACCCCGCTGACCCAAAGGTG 2591
Db 2520 TCGGACCTTCGCCCCCCCCAAGAGTCCAGCCCTTCGCCCCACACCCCGCTGACCCAAAGGTG 2579
Qy 2592 CTGGTAGAAGAGGACACCGCCCGGCTGAGTCTCTGGGGTATCCGTTCCGACGCTGAGAG 2651
Db 2580 CTGGTAGAAGAGGACACCGCCCGGCTGAGTCTCTGGGGTATCCGTTCCGACGCTGAGAG 2639
Qy 2652 GCTTTTCTGGAAAGCCCGCAGCCAGCCCTTCGCGGGGGCTCAGCTCGACTTCTACAGGGC 2711
Db 2640 GCTTTTCTGGAAAGCCCGCAGCCAGCCCTTCGCGGGGGCTCAGCTCGACTTCTACAGGGC 2699
Qy 2712 CACATCACCGCTTCTTCGGGCCAACAAAGGGGCGGCAAGACCAACACCTGTCCTATG 2771
Db 2700 CACATCACCGCTTCTTCGGGCCAACAAAGGGGCGGCAAGACCAACACCTGTCCTATG 2759
Qy 2772 AGTGGCTCTTCCACCCAGTGGTGGCTCTGCTTCATGCTGGGCCACGACGTCCGCTCC 2831
Db 2760 AGTGGCTCTTCCACCCAGTGGTGGCTCTGCTTCATGCTGGGCCACGACGTCCGCTCC 2819
Qy 2832 AGCATGGCGCCATCCGGCCCCACCTGGGCGTCTGCTCAGTACAAAGTGTGTTGAC 2891
Db 2820 AGCATGGCGCCATCCGGCCCCACCTGGGCGTCTGCTCAGTACAAAGTGTGTTGAC 2879
Qy 2892 ATGTGACCGTGGACGAGCACGCTCTGTGTTCTATGGGCGGCTGAAGGCTCTGAGTCCGCT 2951
Db 2880 ATGTGACCGTGGACGAGCACGCTCTGTGTTCTATGGGCGGCTGAAGGCTCTGAGTCCGCT 2939
Qy 2952 GTAGTGGGCCCCGAGAGGACCGCTCTGCTGACAGGATGTGGGCTGGTCTCAAGCAGAGT 3011
Db 2940 GTAGTGGGCCCCGAGAGGACCGCTCTGCTGACAGGATGTGGGCTGGTCTCTCAAGCAGAGT 2999
Qy 3012 GTGCAGACTCGCACCTCTCTGGTGGATGCAACGGAGACTGTCGTGGCCATTCGCTTT 3071
Db 3000 GTGCAGACTCGGCCA CTTCTCTGGTGGATGCAACGGAGCTGTCGTGGCCATTCGCTTT 3059
Qy 3072 GTGGCGGCTCCAAAGTTGTTATCTCGACAGGCTACGGCTGGCGTGATCTCTCTTCC 3131
Db 3060 GTGGCGGCTCCCAAGTTGTTATCTCGACAGGCTACGGCTGGCGTGATCTCTCTTCC 3119
Qy 3132 CGCCCGGTATTTGGGAGTGCTGTCAAAATACGAGAAAGTTCGACGCTGATCTCTCC 3191
Db 3120 CGCCCGGTATTTGGGAGTGCTGTCAAAATACGAGAAAGTTCGACGCTGATCTCTCC 3179
Qy 3192 ACCACCACTGGATGAGGACAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGC 3251
Db 3180 ACCACCACTGGATGAGGACAGCTGTGGGAGACCGTGTGGCGTGGTGGCAGGTGGC 3239
Qy 3252 CGTTGTGCTGTGGCTCCCACTCTTCTCGCGCGCTCACTGGGCTCCGGCTACTAC 3311
Db 3240 CGTTGTGCTGTGGATCCCACTCTTCTCGCGCGCTCACTGGGCTCCGGCTACTAC 3299
Qy 3312 CTGACGCTGGTGAAGCCCGCTGCTGACCAACCAATGAAGAGGCTGACATGACATG 3371
Db 3300 CTGACGCTGGTGAAGCCCGCTGCTGACCAACCAATGAAGAGGCTGACATGACATG 3359
Qy 3372 GAGGCACTGTGGACACCAAGCAGGAAAGAGATGACGAGGAGGAGGAGTGGCTGGC 3431
Db 3360 GAGGCACTGTGGACACCAAGCAGGAAAGAGATGAGGAGGAGGAGTGGCTGGC 3419
Qy 3432 ACTCTCAGCTGTGGCTGTGATCAGCACTGGGTCGCGGGCACGGCTGGTGGAGGAG 3491
Db 3420 ACTCTCAGCTGTGGCTGTGATCAGCACTGGGTCGCGGGCACGGCTGGTGGAGGAG 3479
Qy 3492 CTGCCACAGAGCTGGTGTGCTGCCCTACACGGGTGGCCATGACGGGAGCTTTCGCC 3551
Db 3480 CTGCCACAGAGCTGGTGTGCTGCCCTACACGGGTGGCCATGACGGGAGCTTTCGCC 3539
Qy 3552 ACATCTTCCGAGAGCTAGACCGGGCTGGCGAGCTGAGGCTCACTGGGATC 3611
Db 3540 ACATCTTCCGAGAGCTAGACCGGGCTGGCGAGCTGAGGCTCACTGGGATC 3599

Qy 3612 TCCGACACGAGCTCGAGGAGATCTTCTGAAGGTGGTGGAGAGTGTCTCGGACACA 3671
Db 3600 TCCGACACGAGCTCGAGGAGATCTTCTGAAGGTGGTGGAGAGTGTCTCGGACACA 3659
Qy 3672 GATATGGAGGATGGGAGCTGGGGGAGACCTATGCAACGAGCATTTCTGGCCCTAGACGTA 3731
Db 3660 GATATGGAGGATGGGAGCTGGGGGAGACCTATGCAACGAGCATTTCTGGCCCTAGACGTA 3719
Qy 3732 ACCCTGCGGCTCAAGATCCGCCACAGAGACAGCGCTGGAGAACGGGGAAACAGCTGG 3791
Db 3720 ACCCTGCGGCTCAAGATCCGCCACAGAGACAGCGCTGGAGAACGGGGAAACAGCTGG 3779
Qy 3792 TCAGCCCCAGAGACTGACAGGCTCTGGGGCAGACGCGCTGGGGCCGGGTACAGGGCTGG 3851
Db 3780 TCAGCCCCAGAGACTGACAGGCTCTGGGGCAGACGCGCTGGGGCCGGGTACAGGGCTGG 3839
Qy 3852 GCATGACCCCGCAGCAGCTCGAGGCTCGAGGCTCTCTCAAGCGCTTTCTGGTTCGCCCGC 3911
Db 3840 GCATGACCCCGCAGCAGCTCGAGGCTCTCGAGGCTCTCTCAAGCGCTTTCTGGTTCGCCCGC 3899
Qy 3912 AGCCGCGCGGCTGTTCCGCCAGATGCTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3971
Db 3900 AGCCGCGCGGCTGTTCCGCCAGATGCTGCTGCTGCCCTCTTTGTGGGCTGGCCCTC 3959
Qy 3972 GTGTTTACGCTCATCTGCTGCTCTTTCGGGCACTACCCGGCTCTCGGGCTCAGTCCAC 4031
Db 3960 GTGTTTACGCTCATCTGCTGCTCTTTCGGGCACTACCCGGCTCTCGGGCTCAGTCCAC 4019
Qy 4032 ATGTACGCTGCTCAGGCTGCTCTTCTCAGTGAAGACGCCACAGGGACCTCTGGAAGTGGC 4091
Db 4020 ATGTACGCTGCTCAGGCTGCTCTTCTCAGTGAAGACGCCACAGGGACCTCTGGAAGTGGC 4079
Qy 4092 CGGCTGCTCAGGCGCTGCTCAGAGGACAGACTGGAGGACCCCGAGTGCAGATAGC 4151
Db 4080 CGGCTGCTCAGGCGCTGCTCAGAGGACAGACTGGAGGACCCCGAGTGCAGATAGC 4139
Qy 4152 TCCACAGCTTCTCGGCAACAGAAATCTCTGTAAGTGGCCAAAGTCTTGGCCAGTGGC 4211
Db 4140 TCCACAGCTTCTCGGCAACAGAAATCTCTGTAAGTGGCCAAAGTCTTGGCCAGTGGC 4199
Qy 4212 AACTGGAACCCAGAGTCTCCATCCCGAGCTGCGACGCTGAGTGCAGCCCGGCTGCCGGC 4271
Db 4200 AACTGGAACCCAGAGTCTCCATCCCGAGCTGCGACGCTGAGTGCAGCCCGGCTGCCGGC 4259
Qy 4272 CTGCTGCCGACTGCCCGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4331
Db 4260 CTGCTGCCGACTGCCCGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4319
Qy 4332 TCTGGGAAAGTGGTTCAGAACTTCAGAGCCGGCCGAACTGTCTGACTTCTTGGTCAAGACC 4391
Db 4320 TCTGGGAAAGTGGTTCAGAACTTCAGAGCCGGCCGAACTGTCTGACTTCTTGGTCAAGACC 4379
Qy 4392 TACCGCGCTGCTGGTGGCGCAGGCTGGAAGTGAAGAGTGGGTGAATGAGGTCAAGTAC 4451
Db 4380 TACCGCGCTGCTGGTGGCGCAGGCTGGAAGTGAAGAGTGGGTGAATGAGGTCAAGTAC 4439
Qy 4452 GGAGGCTCTCGCTGGGGGGCCGAGACCCAGGCTGCTGCCCTGGGCGAAGAGTGGGGCGC 4511
Db 4440 GGAGGCTCTCGCTGGGGGGCCGAGACCCAGGCTGCTGCCCTGGGCGAAGAGTGGGGCGC 4499
Qy 4512 TCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCTGCTGCTGGCGGGCCCTCGACCGCTGTC 4571
Db 4500 TCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCTGCTGCTGGCGGGCCCTCGACCGCTGTC 4559
Qy 4572 CTGAAAAACCTCAACAGCTGGGCTCAACGCTGAGTGTCAAGACAGTCTCAAGATCTGG 4631
Db 4560 CTGAAAAACCTCAACAGCTGGGCTCAACGCTGAGTGTCAAGACAGTCTCAAGATCTGG 4619
Qy 4632 TTCAACAAAGAGCTGGCACTCCATGGTGGCTTTGTCAACCGGAGCCAGCAACGCAATC 4691
Db 4620 TTCAACAAAGAGCTGGCACTCCATGGTGGCTTTGTCAACCGGAGCCAGCAACGCAATC 4679

Db	6235	TGCCCCCGCAAGTCCACGCCGCACGGCCTTCGTGGCGGCGAGTTCCCTTGSGTCGG	6294
Qy	6346	AGCTGCGCAGGACAATGAGAGGCGCGTGGCGCTTCAGACTGCCCGCCGGGAGGCGCTGCG	6405
Db	6295	AGTTCGCGAGGACAATGAGAGTTCGCTTCAGCTTCGCCCGGAGGCGGCTGCG	6354
Qy	6406	CCCTGGCGCGCGTCTTTGGAGAGCTGCGCGTGCACGCGCAGAGCACGCGTGGAGGACT	6465
Db	6355	CCCTGGCGCGCGTCTTTGGAGAGCTGCGCGTGCACGCGCGAGAGCACGCGCTGGAGGACT	6414
Qy	6466	TTTTCCGTGAGCCACAGCATGCTGGAGAGATTCTTTGTACTTCTCAAAGAACCGAGGA	6525
Db	6415	TTTTCCGTGAGCCACAGCATGCTGGAGAGATTCTTTGTACTTCTCAAAGAACCGAGGA	6474
Qy	6526	AGGACGAGGACACCGAAGAGCAGAAGAGGACGAGGTGGAGTGGACCCCGCGCCAGGCC	6585
Db	6475	AGGACGAGGACACCGAAGAGCAGNAAGAGCGAGGTGGAGTGGACCCCGCGCCAGGCC	6534
Qy	6586	TGCAGCACCCCAAAACGCGTCAGCGATTCTTCGATGACCCTTAGCACTGCCGAGACTGTGC	6645
Db	6535	TGCAGCACCCCAAAACGCGTCAGCGATTCTTCGATGACCCTTAGCACTGCCGAGACTGTGC	6594
Qy	6646	TCTGAGCCTCCCTCCCTCGGGGCGCGGGGAGGCCCTCGGAATGCAAGGCAAGGTA	6705
Db	6595	TCTGAGCCTCCCTCCCTCGGGGCGCGGGGAGGCCCTCGGAATGCAAGGCAAGGTA	6654
Qy	6706	GAGTGCCTTAGGAGCCCTGGACTCAGGCTGGCAGAGGGCGTGGTGCCCTGGAGAAAAATAA	6765
Db	6655	GAGTGCCTTAGGAGCCCTGGACTCAGGCTGGCAGAGGGCGTGGTGCCCTGGAGAAAAATAA	6714
Qy	6766	GAGNAGCCTGGAGAGACCGTGTCTGTGAAAAAAA	6804
Db	6715	GAGNAGCCTGGAGAGACCGTGTCTGTGAAAAAAA	6753

RESULT 5

ABL57810	ID	ABL57810 standard; DNA; 7795 BP.
XX	AC	
XX	AC	ABL57810;
XX	DT	05-AUG-2002 (first entry)
XX	DT	
XX	DE	Human ABCA7 coding sequence.
XX	XX	
XX	KW	Human; ABCA7; promoter; immunomodulatory; antiinflammatory; metabolic; ATP-binding Cassette; lipid metabolism disorder; immune response; inflammation; gene therapy; gene; ds.
XX	KW	
XX	XX	
XX	OS	Homo sapiens.
XX	XX	
XX	PN	WO200234903-A2.
XX	XX	
XX	PD	02-MAY-2002.
XX	XX	
XX	PF	17-OCT-2001; 2001WO-FR003219.
XX	XX	
XX	PR	24-OCT-2000; 2000FR-00013649.
XX	PR	28-NOV-2000; 2000US-0253141P.
XX	XX	
XX	PA	(AVET) AVENTIS PHARMA SA.
XX	PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX	XX	
XX	PI	Denefle P, Rosier M, Prades C, Arnould-Reguigne I;
XX	PI	Oserio Y Forteau, Duverger N, Chimini G;
XX	XX	
XX	DR	WPI; 2002-362799/39.
XX	XX	
XX	PT	New promoter of the ABCA7 gene, useful for identifying modulators of transcription and in gene therapy of e.g. disorders of lipid metabolism.
XX	PT	
XX	PS	Disclosure; Page 124-125; 126pp; French.

QY 896 ACCTAGACGACAGTGGGGCCCTCCCTCAACTGGTACGAGGCTAGTGACCTGATGAGCT 955
DB 1894 ACCTAGACGACAGTGGGGCCCTCCCTCAACTGGTACGAGGCTAGTGACCTGATGAGCT 1953
QY 956 GGTGGGCGAGGACGAGATCCGCTGCGACAGCAGCCTGAGCCCGCTGCTCGGA 1015
DB 1954 GGTGGGCGAGGACGAGATCCGCTGCGACAGCAGCCTGAGCCCGCTGCTCGGA 2013
QY 1016 GCTGATTGGAGCCCTGGGACAGCACCCGCTGTCGCCCTGCTCTGGAGACGCTGAAGCC 1075
DB 2014 GCTGATTGGAGCCCTGGGACAGCACCCGCTGTCGCCCTGCTCTGGAGACGCTGAAGCC 2073
QY 1076 TCTGATCTCTGGGAAGCTACTCTTTGCAACAGATACACTTTTACCCGGAAGCTCATGBC 1135
DB 2074 TCTGATCTCTGGGAAGCTACTCTTTGCAACAGATACACTTTTACCCGGAAGCTCATGBC 2133
QY 1136 CCAGGTGAACCGGACCTTCGAGGAGCTCACCTCTGTAGGAGTGTCCGGAGGTGGGA 1195
DB 2134 CCAGGTGAACCGGACCTTCGAGGAGCTCACCTCTGTAGGAGTGTCCGGAGGTGGGA 2193
QY 1196 GATGCTGGGACCCCGATCTTCACCTTCATGAACGACAGTTCCAATGTGCCATGCTGCA 1255
DB 2194 GATGCTGGGACCCCGATCTTCACCTTCATGAACGACAGTTCCAATGTGCCATGCTGCA 2253
QY 1256 GCGGCTCTGTCAGATGTCAGGATGAAGGAAGAGGACGCCAGACCTGGAGGCCGGGACCA 1315
DB 2254 GCGGCTCTGTCAGATGTCAGGATGAAGGAAGAGGACGCCAGACCTGGAGGCCGGGACCA 2313
QY 1316 CATGAGAGCCCTGCCATCTTCTGGAAGCTGCGGACGCTGGGAGCGGTGCTACAGTGGCAGGACGC 1375
DB 2314 CATGAGAGCCCTGCCATCTTCTGAGCCCTGGGAGCGGTGCTACAGTGGCAGGACGC 2373
QY 1376 ACACGCTGATGTGGGGCACTGGTGGGACGCTGGGCGAGTGACGAGTGCTGCTCTT 1435
DB 2374 ACACGCTGATGTGGGGCACTGGTGGGACGCTGGGCGAGTGA CCGAGTGCTGCTCTT 2433
QY 1436 GGAAGCTGGAGGCGGACCCCTCAGAGGACGCCCTGGTGTGCGGGCCCTGCAACTGCT 1495
DB 2434 GGAAGCTGGAGGCGGACCCCTCAGAGGACGCCCTGGTGTGCGGGCCCTGCAACTGCT 2493
QY 1496 GCGGAAATCGATTCGGGCGGCGTCTTCTTTGGGACCTGAGGACCTTTCAGACCC 1555
DB 2494 GCGGAAATCGATTCGGGCGGCGTCTTCTTTGGGACCTGAGGACCTTTCAGACCC 2553
QY 1556 CACAGACACCAACCCAGACCTGGGCGCCCGGACCGTGGCATCAAAATCCGCATGGA 1615
DB 2554 CACAGACACCAACCCAGACCTGGGCGCCCGGACCGTGGCATCAAAATCCGCATGGA 2613
QY 1616 CATGACGTGGTACAGGACCAATAAGATCAGGACAGGTTTGGGACCCCTGGCCAGC 1675
DB 2614 CATGACGTGGTACAGGACCAATAAGATCAGGACAGGTTTGGGACCCCTGGCCAGC 2673
QY 1676 GCGGACCCCTGACCGACCTGCGCTACGTGTGGGGCGGCTTGTGTACCTGCAAGACCT 1735
DB 2674 GCGGACCCCTGACCGACCTGCGCTACGTGTGGGGCGGCTTGTGTACCTGCAAGACCT 2733
QY 1736 GGTGAGAGTGCAGCCGTCGCGTGTCTCAGCGGCGCCAAACCCCGGGCGGCGCTTACCT 1795
DB 2734 GGTGAGAGTGCAGCCGTCGCGTGTCTCAGCGGCGCCAAACCCCGGGCGGCGCTTACCT 2793
QY 1796 GCAGCAGATGCCCTATCCGTGTATGTGGAACGAGCTGTCTGCTGTGTGCTGAGCCGGTC 1855
DB 2794 GCAGCAGATGCCCTATCCGTGTATGTGGAACGAGCTGTCTGCTGTGTGCTGAGCCGGTC 2853
QY 1856 GCTCCGCTCTTCTGACGCTGGCTGGATCTACTCCGTGACACTGACAGTGAAGGCGGT 1915
DB 2854 GCTCCGCTCTTCTGACGCTGGCTGGATCTACTCCGTGACACTGACAGTGAAGGCGGT 2913
QY 1916 GGTGGGAGAGGAGACCGGCTGGGGACACCAATGCGCGGCTAGGGGCTCAGCCGCGC 1975
DB 2914 GGTGGGAGAGGAGACCGGCTGGGGACACCAATGCGCGGCTAGGGGCTCAGCCGCGC 2973

QY 1976 GGTGCTCTGGCTAGGCTGGTTCTCTCAGCTGCTCGGGCCCTTCTTCTGCTCAGCGCCGCGCT 2035
DB 2974 GGTGCTCTGGCTAGGCTGGTTCTCTCAGCTGCTCGGGCCCTTCTTCTGCTCAGCGCCGCACT 3033
QY 2036 GCTGGTTCTGGTCTCAAGCTGGGGGACATCTCTCCCTACAGCCACCCGGGCGTGGTCTT 2095
DB 3034 GCTGGTTCTGGTCTCAAGCTGGGGGACATCTCTCCCTACAGCCACCCGGGCGTGGTCTT 3093
QY 2096 CTTGTTCTTGGCAGCTTTCGGGTCGCGGTCAGCTGACCCAGAGCTTCTTCTGCTCAGCGCTT 2155
DB 3094 CTTGTTCTTGGCAGCTTTCGGGTCGCGGTCGCGTGAACCCAGAGCTTCTTCTGCTCAGCGCTT 3153
QY 2156 CTTTCTCCGCGCAACTGGTGTGGCTGTGGCGGCTGTGGGCTACTTCTTCTTCTTACCT 2215
DB 3154 CTTTCTCCGCGCAACTGGTGTGGCTGTGGCGGCTGTGGGCTACTTCTTCTTCTTACCT 3213
QY 2216 GCGCTTACCTGTGTGTGGCTGTGGCGGACCGGCTGCGCGGGTGGCGGCTGGCGCGC 2275
DB 3214 GCGCTTACCTGTGTGTGGCTGTGGCGGACCGGCTGCGCGGGTGGCGGCTGGCGCGC 3273
QY 2276 GAGCCTGTCTGCGCCGCTGGCTTTCGGCTTTCGGCTGCGAGAGCTTGGCTTCTGCTGAGGA 2335
DB 3274 GAGCCTGTCTGCGCCGCTGGCTTTCGGCTTTCGGCTGCGAGAGCTTGGCTTCTGCTGAGGA 3333
QY 2336 GCAAGGCGAGGCGCGCAGTGGCAACAGTGGGCAACCGGCTTACCGGAGAGCTTCTCAG 2395
DB 3334 GCAAGGCGAGGCGCGCAGTGGCAACAGTGGGCAACCGGCTTACCGGAGAGCTTCTCAG 3393
QY 2396 CTTGGCCGAGTCTTCTGGGCTTTCGGCTTTCGGCTGCGAGAGCTTACGGGCTTGGCCACCTG 2455
DB 3394 CTTGGCCGAGTCTTCTGGGCTTTCGGCTTTCGGCTGCGAGAGCTTACGGGCTTGGCCACCTG 3453
QY 2456 GTACTGTGAAGCTGTGTGCCAGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 2515
DB 3454 GTACTGTGAAGCTGTGTGCCAGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 3513
QY 2516 TCGAGGAGCTACTTGGTGGGACCTTGGCCCGGCAAGAGTCCAGCCCTTGGCCCGCC 2575
DB 3514 TCGAGGAGCTACTTGGTGGGACCTTGGCCCGGCAAGAGTCCAGCCCTTGGCCCGCC 3573
QY 2576 GCTGGAACCAAGGTGTGTGAGAGGCAACCGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCT 2635
DB 3574 GCTGGAACCAAGGTGTGTGAGAGGCAACCGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCT 3633
QY 2636 TCGCAGCCTGAGAGGCTTTCCTGGAAGCCCGCAGCAGCCCTTGGCGGGGCTCAGCT 2695
DB 3634 TCGCAGCCTGAGAGGCTTTCCTGGAAGCCCGCAGCAGCCCTTGGCGGGGCTCAGCT 3693
QY 2696 GGACTTCTACAGGGCCACATCAGCGCTTCTTGGGCAACCGGGGCGGCAAGCCAC 2755
DB 3694 GGACTTCTACAGGGCCACATCAGCGCTTCTTGGGCAACCGGGGCGGCAAGCCAC 3753
QY 2756 CACCTGTCTCATCTTGTGTGGCTTTCCTCAACAGTGGTGGCTTGTGCTTCTTCTGCTGGG 2815
DB 3754 CACCTGTCTCATCTTGTGTGGCTTTCCTCAACAGTGGTGGCTTGTGCTTCTTCTGCTGGG 3813
QY 2816 CCAGAGCTCGCTCCAGCATGGCGCCATCCGCGCCCTTGGGCTTGTGCTTCTTCTGCTGCTG 2875
DB 3814 CCAGAGCTCGCTCCAGCATGGCGCCATCCGCGCCCTTGGGCTTGTGCTTCTTCTGCTGCTG 3873
QY 2876 CAAAGTGTGTTTGAACATGCTGACCGTGGAGCAGCAGCTTGTGTTCTATGGGCGGCTGAA 2935
DB 3874 CAAAGTGTGTTTGAACATGCTGACCGTGGAGCAGCAGCTTGTGTTCTATGGGCGGCTGAA 3933
QY 2936 GGGTCTGAGTGCCTGTGTAGTGGCCCGGAGCAGGACCGTGTGCTGCTGAGGATGGGGCT 2995
DB 3934 GGGTCTGAGTGCCTGTGTAGTGGCCCGGAGCAGGACCGTGTGCTGCTGAGGATGGGGCT 3993
QY 2996 GGTCTTCAAGCAGAGTGTGACAGCTCGCCACTTCTTGGTGGGATGCAACGGAAGCTGTC 3055
DB 3994 GGTCTTCAAGCAGAGTGTGACAGCTCGCCACTTCTTGGTGGGATGCAACGGAAGCTGTC 4053
QY 3056 CGTGGCCATTGCTTGTGGGCGGCTCCAAAGTTGTTATCTCTGAGAGGCTTACGGCTGG 3115

Db 4054 CGTGGCAATGCTTTGTGGCGGCTCCCAAGTTGTTATCTTGCAGCGCTACCGCTGG 4113
QY 3116 CGTGGATCTGCTTCCCGCGCGGATTTTGGAGCTGTGCTCAATAATCGGAGAGGTGG 3175
Db 4114 CGTGGATCTGCTTCCCGCGCGGATTTTGGAGCTGTGCTCAATAATCGGAGAGGTGG 4173
QY 3176 CAGCTGATCTCTCCACCCACACCTGGATGAGGAGAGCTGCTGGGAGACCGTGTGGC 3235
Db 4174 CACGCTGATCTCTCCACCCACACCTGGATGAGGAGAGCTGCTGGGAGACCGTGTGGC 4233
QY 3236 TGTGGTGGAGGTGGCGCTTGTGCTGTGCTGCCCTCCCACTCTCTCTGCGCGCTCACCT 3295
Db 4234 CGTGGTGGAGGTGGCGCTTGTGCTGTGCTGCCCTCCCACTCTCTCTGCGCGCTCACCT 4293
QY 3296 GGGCTCCGGCTACTACTGACGCTGTGTGAAGGCCCGCTGCCCTGACCAACCAATGAGAA 3355
Db 4294 GGGCTCCGGCTACTACTGACGCTGTGTGAAGGCCCGCTGCCCTGACCAACCAATGAGAA 4353
QY 3356 GGTGACACTGACATGGAGGGAGTGTGGACACAGCGAGGAAAAGAGATGGCAGCCA 3415
Db 4354 GGTGACACTGACATGGAGGGAGTGTGGACACAGCGAGGAAAAGAGATGGCAGCCA 4413
QY 3416 GGGCAGCAGAGTGGCACTCTCAGCTGTGCTGGCCCTGGTACAGCACTGGGTGCCCGGGC 3475
Db 4414 GGGCAGCAGAGTGGCACTCTCAGCTGTGCTGGCCCTGGTACAGCACTGGGTGCCCGGGC 4473
QY 3476 ACGGCTGGTGGAGAGTGGCACACAGCAGTGTGCTGGTGTGCTGCCCTACACGGGTGCCCA 3535
Db 4474 ACGGCTGGTGGAGAGTGGCACACAGCAGTGTGCTGGTGTGCTGCCCTACACGGGTGCCCA 4533
QY 3536 TGAAGGAGCTTCCGACACTCTTCGAGAGCTAGACACGGCGCTGGCGAGCTGAGGCT 3595
Db 4534 TGAAGGAGCTTCCGACACTCTTCGAGAGCTAGACACGGCGCTGGCGAGCTGAGGCT 4593
QY 3596 CACTGGCTACGGGATCTCCGACACAGCCTCGAGGAGATCTTCTGAAAGTGTGGAGGA 3655
Db 4594 CACTGGCTACGGGATCTCCGACACAGCCTCGAGGAGATCTTCTGAAAGTGTGGAGGA 4653
QY 3656 GTGTGCTGGGACACAGATATGAGGATGGCAGCTCGGGCAGCACCTATGCAAGGCAT 3715
Db 4654 GTGTGCTGGGACACAGATATGAGGATGGCAGCTCGGGCAGCACCTATGCAAGGCAT 4713
QY 3716 TGTGCGCTAGAGTAACTCGGCTCAAGATGTCGCGCAGAGAGACGCGCTGGAGAA 3775
Db 4714 TGTGCGCTAGAGTAACTCGGCTCAAGATGTCGCGCAGAGAGACGCGCTGGAGAA 4773
QY 3776 CGGGGAACAGCTGGGTGAGCCCGAGAGACTGACCGGGCTCTGGGCGAGACGCGTGGG 3835
Db 4774 CGGGGAACAGCTGGGTGAGCCCGAGAGACTGACCGGGCTCTGGGCGAGACGCGTGGG 4833
QY 3836 CGGGGTACAGGGCTGGGCACTGACCCGCGCAGAGCTCGAGGCTCGCTTCTCAAGCGCTT 3895
Db 4834 CGGGGTACAGGGCTGGGCACTGACCCGCGCAGAGCTCGAGGCTCTGCTTCTCAAGCGCTT 4893
QY 3896 TCTGCTTCCCGCGAGCGCGCGGCTGTGTCGCCCAAGATCGTGTGCTGCTGCCCTCTT 3955
Db 4894 TCTGCTTCCCGCGAGCGCGCGGCTGTGTCGCCCAAGATCGTGTGCTGCTGCCCTCTT 4953
QY 3956 TGTGGGCTGGGCTCTGCTGTTACGCTCATCTGTCGCTCTTTCGGGCACTACCGGCTCT 4015
Db 4954 TGTGGGCTGGGCTCTGCTGTTACGCTCATCTGTCGCTCTTTCGGGCACTACCGGCTCT 5013
QY 4016 GCGGCTCAGTCCCACCATGTACCGTGTCTCAGGTGTCTCTTCTCAGTGAGGACGCCCCAGG 4075
Db 5014 GCGGCTCAGTCCCACCATGTACCGTGTCTCAGGTGTCTCTTCTCAGTGAGGACGCCCCAGG 5073
QY 4076 GGAACCTTGAAGTGGTCCCGGCTGTCTGAGGCGCTGTCTGCAAGGAGGAGGAGGAGCC 4135
Db 5074 GGAACCTTGAAGTGGTCCCGGCTGTCTGAGGCGCTGTCTGCAAGGAGGAGGAGGAGCC 5133
QY 4136 CCGAGTGCACATAGCTCCACAGGTTCTCGGACACAGAGTTCTGCTCAAGTGGCCAA 4195

Db 5134 CCAGTGCAGCATAGTCTCCACAGGTTCTCGGCACACAGAAATCTCTGTGAAGTGGCCAA 5193
QY 4196 GGTCTTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCGAGGTGAGCCA 4255
Db 5194 GGTCTTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCGAGGTGAGCCG 5253
QY 4256 GCCCGTGGCCCGGCGCTGCTGCCAGCTGCGCCGCTGAGCTGGTGGTCCCGCTCCGCGC 4315
Db 5254 GCCCGTGGCCCGGCGCTGCTGCCAGCTGCGCCGCTGAGCTGGTGGTCCCGCTCCGCGC 5313
QY 4316 CCAGGCAGTGCACCGGCTCTGGGGAAGTGTTCAGAACTTGACAGGCGCGAACTGTCTGA 4375
Db 5314 CCAGGCAGTGCACCGGCTCTGGGGAAGTGTTCAGAACTTGACAGGCGCGAACTGTCTGA 5373
QY 4376 CTTCTGTGTCAAGACCTACCCGCGCTGGTGGCCAGGCGCTGGAAGACTTAAAGAGTGGGT 4435
Db 5374 CTTCTGTGTCAAGACCTACCCGCGCTGGTGGCCAGGCGCTTAAAGACTTAAAGAGTGGGT 5433
QY 4436 GAATGAGGTCAAGTACCGAGGCTTCTGCTGGGGGCCGAGACCCAGGCTTGCCTCGGG 4495
Db 5434 GAATGAGGTCAAGTACCGAGGCTTCTGCTGGGGGCCGAGACCCAGGCTTGCCTCGGG 5493
QY 4496 CCAAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGTCCCTGCTGGCGG 4555
Db 5494 CCAAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGTCCCTGCTGGCGG 5553
QY 4556 GGGCTTGCACCGTGTCTTGAATAAACTCAAGCTGGGCTCAAGGCTTGGATGCTCAGGA 4615
Db 5554 GGGCTTGCACCGTGTCTTGAATAAACTCAAGCTGGGCTCAAGGCTTGGATGCTCAGGA 5613
QY 4616 CAGTCTCAAGATCTGTTTCAACAAAGAGCTGGCACTCCATGTTGGCTTGTGTCAACCG 4675
Db 5614 CAGTCTCAAGATCTGTTTCAACAAAGAGCTGGCACTCCATGTTGGCTTGTGTCAACCG 5673
QY 4676 AGCAGAAACGCAATCTCTCGTCTCACCTGCCCGCCAGCGCGCGCCAGCCACAG 4735
Db 5674 AGCAGAAACGCAATCTCTCGTCTCACCTGCCCGCCAGCGCGCGCCAGCCACAG 5733
QY 4736 CATCACCACTCAACACCCCTTGAACTCAACAAAGAGAGAGCTGTCTGAGGCTGCACT 4795
Db 5734 CATCACCACTCAACACCCCTTGAACTCAACAAAGAGAGAGCTGTCTGAGGCTGCACT 5793
QY 4796 GATGGCTCTCTCGGTGGAGCTCTGCTCCATCTGTGTGGTCTTGTGGCATGTCTTGT 4855
Db 5794 GATGGCTCTCTCGGTGGAGCTCTGCTCCATCTGTGTGGTCTTGTGGCATGTCTTGT 5853
QY 4856 CCGGCGAGCTTCACTCTTGTCTCATTTGAGGAGCGAGTCAACCGAGCAAGCACCTGCA 4915
Db 5854 CCGGCGAGCTTCACTCTTGTCTCATTTGAGGAGCGAGTCAACCGAGCAAGCACCTGCA 5913
QY 4916 GCTCATGGGGGCGCTGTCCCACTCTGCTTACTGGCTTGGCAACTTCTCTGGGACATGTG 4975
Db 5914 GCTCATGGGGGCGCTGTCCCACTCTGCTTACTGGCTTGGCAACTTCTCTGGGACATGTG 5973
QY 4976 TAACTACTTGTGTGCAGCATGATGTGTGTCTCATCTTCTGGCGCTTCCAGCAGAGGGC 5035
Db 5974 TAACTACTTGTGTGCAGCATGATGTGTGTCTCATCTTCTGGCGCTTCCAGCAGAGGGC 6033
QY 5036 ATATGTGGCGCTTCCCAACTGCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5095
Db 6034 ATATGTGGCGCTTCCCAACTGCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 6093
QY 5096 GATCACACGCTCATGTATCCAGGCTCTCTTCTTCTCTGCTGCCAGCAGCAGCTATGT 5155
Db 6094 GATCACACGCTCATGTATCCAGGCTCTCTTCTTCTCTGCTGCCAGCAGCAGCTATGT 6153
QY 5156 GGTGCTCAGCTGCATAACTCTTTATTGGCATCAATGAAGCATGGCCACCTTTGTGCT 5215
Db 6154 GGTGCTCAGCTGCATAACTCTTTATTGGCATCAATGAAGCATGGCCACCTTTGTGCT 6213
QY 5216 TGAAGTCTTCTGTGATCAGAAAGCTGCAGGAGGTGAGCCGGATCTTGAAGACAGGTCTTCT 5275
Db 6214 TGAAGTCTTCTGTGATCAGAAAGCTGCAGGAGGTGAGCCCGATCTTGAAGACAGGTCTTCT 6273

Qy	870	TCAGAGGCCCTCTGCAGTGTTCAGGGGACCTTAGCAGCACAGATGGGGCCCTCCCTCAACTGG	929
Db	655	TCAGAGGCCCTCTGCAGTGTTCAGGGGACCTTAGCAGCACAGATGGGGCCCTCCCTCAACTGG	714
Qy	930	TACAGAGCTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGANAATCGGCCCTGCCAGAC	989
Db	715	TACAGAGCTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGANAATCGGCCCTGCCAGAC	774
Qy	990	AGCAGCCTGAGCCCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCGCTGTCC	1049
Db	775	AGCAGCCTGAGCCCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCGCTGTCC	834
Qy	1050	CGCCTGCTCTGGAGACGCCTGAAGCCCTCTGATCTCTGGGAGGTACTCTTTTGACACCAAT	1109
Db	835	CGCCTGCTCTGGAGACGCCTGAAGCCCTCTGATCTCTGGGAGGTACTCTTTTGACACCAAT	894
Qy	1110	ACACCTTTTACCCCGGAAGCTCATGGCCAGAGTGAAACCGGACCTTTCGAGGAGCTCACCCCTG	1169
Db	895	ACACCTTTTACCCCGGAAGCTCATGGCCAGAGTGAAACCGGACCTTTCGAGGAGCTCACCCCTG	954
Qy	1170	CTGAGGGATGTCGSGGAGGTGTGGGAGATGCTGSGGACCCCGGATCTTCACTTCATGAAC	1229
Db	955	CTGAGGGATGTCGSGGAGGTGTGGGAGATGCTGSGGACCCCGGATCTTCACTTCATGAAC	1014
Qy	1230	GACAGTTTCCAAATGTGGCCA TGCTGCAGCGGCTCTTCAGATGACGAGTGAAGGAAGAAGG	1289
Db	1015	GACAGTTTCCAAATGTGGCCA TGCTGCAGCGGCTCTTCAGATGACGAGTGAAGGAAGAAGG	1074
Qy	1290	CAGCCAGACCTTGAGAGCGGGAGACCAATGAGAGCCCTGCGATCTTTCTGGACCCCTGGG	1349
Db	1075	CAGCCAGACCTTGAGAGCGGGAGACCAATGAGAGCCCTGCGATCTTTCTGGACCCCTGGG	1134
Qy	1350	AGCGGTGCTACAGCTGCGAGGACGACAGCTGATGTGGGGACCTGTGTGGGACGCGCTG	1409
Db	1135	AGCGGTGCTACAGCTGCGAGGACGACAGCTGATGTGGGGACCTGTGTGGGACGCGCTG	1194
Qy	1410	GGCCGAGTGACGGAAGTGCTGTCTTGGAACAAGCTGGAGGCGGCACCCCTCAGAGGCGAGCC	1469
Db	1195	GGCCGAGTGACGGAAGTGCTGTCTTGGAACAAGCTGGAGGCGGCACCCCTCAGAGGCGAGCC	1254
Qy	1470	CTGTGTTCGGGGCCCTGCAACTCTCTCGGGAA CATCGATTCTGGGGCGGGGTCTTCTTC	1529
Db	1255	CTGTGTTCGGGGCCCTGCAACTCTCTCGGGAA CATCGATTCTGGGGCGGGGTCTTCTTTC	1314
Qy	1530	TTGGGACCTCAGGACTCTTCAGACCCACAGAGCACCCAACCCACAGACCTTGGGCCCGGC	1589
Db	1315	TTGGGACCTCAGGACTCTTCAGACCCACAGAGCACCCAACCCACAGACCTTGGGCCCGGC	1374
Qy	1590	CACGTGCGCATCAAAATCCGATCGGACATTTGACGTGTTCACGAGGACCAATAGATCAGG	1649
Db	1375	CACGTGCGCATCAAAATCCGATCGGACATTTGACGTGTTCACGAGGACCAATAGATCAGG	1434
Qy	1650	GACAGTTTGGGACCTTGCGCCAGCCGCGAACCCCTTGACCGACCTTGCCTGCTAGTGTGG	1709
Db	1435	GACAGTTTGGGACCTTGCGCCAGCCGCGAACCCCTTGACCGACCTTGCCTGCTAGTGTGG	1494
Qy	1710	GGCGGCTTCGTGTACCTTGCAAGACCTTGGTGGAGCGTGCAGCGCTCCGCTGCTCAGCGGC	1769
Db	1495	GGCGGCTTCGTGTACCTTGCAAGACCTTGGTGGAGCGTGCAGCGCTCCGCTGCTCAGCGGC	1554
Qy	1770	GCCAAACCCCGGGCGGGCTCTACCTGACAGATGCCCTATCTCGTGCTATTTGGAGCGAC	1829
Db	1555	GCCAAACCCCGGGCGGGCTCTACCTGACAGATGCCCTATCTCGTGCTATTTGGAGCGAC	1614
Qy	1830	GTGTTCCTGTGTGCTGAGCGCGTGTGCGCTCTTCTCTGACGCTTGGGCTGTGATCTAC	1889
Db	1615	GTGTTCCTGTGTGCTGAGCGCGTGTGCGCTCTTCTCTGACGCTTGGGCTGTGATCTAC	1674
Qy	1890	TCCGTGACATGACAGTGAGGCGGTGTGCGGAGAGGAGACCGCGCTTCCGGGACACC	1949
Db	1675	TCCGTGACATGACAGTGAGGCGGTGTGCGGAGAGGAGACCGCGCTTCCGGGACACC	1734
Qy	1950	ATGCGCGCATGGGGCTCAGCCGCGCGGTGTCTCTGGCTAGGCTGGTTTCTTCAGCTGCTC	2009

1735	A T G C G C G C A T G G G G C T C A G C C G C G G T G C T C T G C G T A G G C T G G T T C C T C A G T G C T C	1794
2010	G G G C C T T C T G C T C A G C G C G C G T G C T T C T G T T C T G T G C T C A A G C T G G G G A C A T C T C	2069
1795	G G G C C T T C T G C T C A G C G C G C G T G C T T C T G T T C T G T G C T C A G C T G G G G A C A T C T C	1854
2070	C C T A C A G C C A C C G G G C T G T C T T C T G T T C T T T G C A G C C T C G C G T G C C A C G T G	2129
1855	C C T A C A G C C A C C G G G C T G T C T T C T G T T C T T G G C A G C C T C C G G T G C C C A C G T G	1914
2130	A C C A A G C T T C T G C T C A G G C C T T C T C T C G G G C C A A C C T G G C T G C G C C T G C G G C	2189
1915	A C C A A G C T T C T G C T C A G G C C T T C T C T C C G C G C C A A C C T G G C T G C G C C T G C G G C	1974
2190	G G C T G G C C T A C T T C C C T C T A C C T G C C C T A G C T G C T G T G T G G C T T G C G G A C C G G	2249
1975	G G C C T G G C C T A C T T C C C C T C T A C C T G C C C T A G C T G T G T G T G G C T T T G C G G A C C G G	2034
2250	C T G C C C G G G T G G C C G C T G G C G C A G C T G C T G T G C C C G T G G C C T C G G C T T C G C G C	2309
2035	C T G C C C G G G T G G C C G C T G G C G C A G C C T G C T G C C C G T G G C C T C G G C T T C G C C	2094
2310	T G C A G A G C C T G G C T C T G C T G A G A G C A A G G G C A G G G C G C A G T G C A C A A C G T G G C C	2369
2095	T G C A G A G C C T G G C T C T G C T G A G A G C A A G G G C A G G G C G C A G T G G C A C A A C G T G G C C	2154
2370	A C C G G C C T A C G G C A G C G T T C A G C T G G C C C A G E T C T C T G G C C T T C T G C T G C T G C A C	2429
2155	A C C G G C C T A C G G C A G C G T T C A G C T G G C C C A G E T C T C T G G C C T T C T G C T G C T G C A C	2214
2430	G G G C G C C T A C G G C C T C G C C A C C T G T A C C T G G A A G C T G T G T G C C A G G C A G T A C G G G	2489
2215	G G G C G C C T A C G G C C T C G C C A C C T G T A C C T G G A A G C T G T G T G C C A G G C C A G T A C G G	2274
2490	A T C C C T G A A C C A T G G A A T T T T C T T T T C G G A G A G C T A C T G T G C G A C C T T C G G C C C C C C	2549
2275	A T C C C T G A A C C A T G G A A T T T T C T T T T C G G A G A G C T A C T G T G C G A C C T T C G G C C C C C C	2334
2550	A A G A G T C A G C C C C T T C C C C C C C C C G C T G A C C A A A G G T G C T G G T A G A A G G C A C C G	2609
2335	A A G A G T C A G C C C C T T G C C C C A C C C G C T G A C C A A A G G T G C T G G T A G A A G A G C A C C G	2394
2610	C C C G C C T G A G T C C T G C G T A T C C G T T C G C A G C T G G A A A G C G C T T T C T G A A G C C C G	2669
2395	C C C G C C T G A G T C C T G G G T C T C G T T C G C A G C C T G G A A G C G C T T T C T G A A G C C C G	2454
2670	C A G C A G C C C T G C G G G G C T A G C C T G A C T T C T A C A G G G C C A C A C C C C C T T C T G	2729
2455	C A G C A G C C C T G C G G G G C T C A G C C T G A C T T C T A C A G G G C C A C A C C C C C T T C T G	2514
2730	G G C C A C A A G G G G C G G C A G A C C A C C C T G C C A C T T T G A G T G G C C T C T C C C A C C C	2789
2515	G G C C A C A G C G G G C C G C A A G C C A C C C T G T C C A T C T T G A G T G G C C T C T T C C C A C C C	2574
2790	A G T G T G C C T G C C T C A C C T C T G G C C A G A G C C G C C C A C A T G G C C G C C A T C C G G	2849
2575	A G T G T G C C T G C C T C A C C T C T G G C C A G A G C C G C C C A C A T G C A T G C C G C C A T C C G G	2634
2850	C C C C A C C T G G G C G T C T G C T C C T A G T A C A A C G T G C T T T T G A C A T G C T A C C G T G G A C G A G	2909
2635	C C C C A C C T G G G C G T C T G C T C C T A G T A C A A C G T G C T T T T G A C A T G C T A C C G T G G A C G A G	2694
2910	C A G C T C T G G T C T A T G G G C G G C T G A A G G T C T G A G T G C G C T G A T G G G C C C C G A C G A G	2969
2695	C A G C T C T G G T C T A T G G G C G G C T G A A G G G T C T G A G T G C G C T G T G G T G G G C C C C G A C G A G	2754
2970	G A C C G T C T G C A G A T G T G G G C T G T C T C C A A G C A G A G T G T G A C A C T C G C C A C C T C	3029
2755	G A C C G T C T A C T G A G A T G T G G G C T G G T C T C C A A G C A G A G T G T C A C A C T C G C C A C C T C	2814
3030	T C T G T G G G A T G C A A C G A A G C T G T C C G T G G C C A T T G C C T T T G G G G C G T C C C A A G T T	3089

Db 2815 TCTGTGGATGCAACGGAAGCTGTCTGCTGGCCATTGCTTTGTGGCGGCTCCCAAGTT 2874
Qy 3090 GTTATCTTGGAGAGCCTACGGCTGGCGGTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG 3149
Db 2875 GTTATCTTGGAGAGCCTACGGCTGGCGGTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG 2934
Qy 3150 GTGCTGCTCAATATACCGAAGAGGTGGACGCTGATCTCTCAACCCACCACTTGATGAG 3209
Db 2935 GTGCTGCTCAATATACCGAAGAGGTGGACGCTGATCTCTCAACCCACCACTTGATGAG 2994
Qy 3210 GCAGAGCTGCTGGGAGACCGTGTGCTGTGTGGCAGGTGGCGCTTGTGCTGTGCTGGC 3269
Db 2995 GCAGAGCTGCTGGGAGACCGTGTGCGCTGTGTGGCAGGTGGCGCTTGTGCTGTGCTGGC 3054
Qy 3270 TCCCACTCTTCTCGCGCGCTCACTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC 3329
Db 3055 TCCCACTCTTCTCGCGCGCTCACTGGGCTCCGGCTACTACTGACGCTGTGTAAGGCC 3114
Qy 3330 GCGTGCCTTGACCAACCAATGAGAAGGCTGACACTGACATGAGAGGCGAGTGTGACACC 3389
Db 3115 GCGTGCCTTGACCAACCAATGAGAAGGCTGACACTGACATGAGAGGCGAGTGTGACACC 3174
Qy 3390 AGGCAGGAAAGAAAGAAATGGCAGCCAGGCGACAGAGTCCGACCTCTCAGCTGTGGCC 3449
Db 3175 AGGCAGGAAAGAAAGAAATGGCAGCCAGGCGACAGAGTCCGACCTCTCAGCTGTGGCC 3234
Qy 3450 CTGTGTACAGCACTGGGTGCGCGGCGACCGGTGTGTGGAGAGCTGCCACACGAGCTGTGTG 3509
Db 3235 CTGTGTACAGCACTGGGTGCGCGGCGACCGGTGTGTGGAGAGCTGCCACACGAGCTGTGTG 3294
Qy 3510 CTGTGTGCTGCCCTACACGGGTGCCCATGACCGGAGCTTTCGCACTCTTTCCGAGAGCTA 3569
Db 3295 CTGTGTGCTGCCCTACACGGGTGCCCATGACCGGAGCTTTCGCACTCTTTCCGAGAGCTA 3354
Qy 3570 GACACGCGCTGGCGGAGCTGAGGCTCTCTGCTACGGGATCTCCGACACAGACCTCGAG 3629
Db 3355 GACACGCGCTGGCGGAGCTGAGGCTCTCTGCTACGGGATCTCCGACACAGACCTCGAG 3414
Qy 3630 GAGATCTTCTGAGGTGTGGAGAGTGTGTGGGACACAGATATGAGAGATGGCAGC 3689
Db 3415 GAGATCTTCTGAGGTGTGGAGAGTGTGTGGGACACAGATATGAGAGATGGCAGC 3474
Qy 3690 TGGGCGACGACCTATGACAGGCAATGCTGGCTTAGACGTAACCTCGGCTCAAGATG 3749
Db 3475 TGGGCGACGACCTATGACAGGCAATGCTGGCTTAGACGTAACCTCGGCTCAAGATG 3534
Qy 3750 CCGCCACAGGAGACAGCGCTGGAGAACCGGGGACACAGCTGGGTGAGCCCCAGACTGAC 3809
Db 3535 CCGCCACAGGAGACAGCGCTGGAGAACCGGGGACACAGCTGGGTGAGCCCCAGACTGAC 3594
Qy 3810 CAGGCTCTGGGCCAGACCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3869
Db 3595 CAGGCTCTGGGCCAGACCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCCGCAGCAG 3654
Qy 3870 CTCACGGCCCTGCTTCTCAAGCGCTTCTGCTGTGCGCGCAGCGCGCGGCGCTGTTC 3929
Db 3655 CTCACGGCCCTGCTTCTCAAGCGCTTCTGCTGTGCGCGCAGCGCGCGGCGCTGTTC 3714
Qy 3930 GCGCAGATGCTGCTGCTGCTCTTGTGTGGGCTGGGCGCTGCTGCTGCTGCTGCTGCTG 3989
Db 3715 GCGCAGATGCTGCTGCTGCTCTTGTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3774
Qy 3990 CCTCTTTCCGGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGCTGCTCAGGTG 4049
Db 3775 CCTCTTTCCGGGCACTACCGGCTCTGGGCTCAGTCCACCATGTACGCTGCTCAGGTG 3834
Qy 4050 TCCTTTCTTCAGTGAGGACGCCCCAGGGGACCTTGACCGTGGCCGCTGCTCGAGCGCTG 4109
Db 3835 TCCTTTCTTCAGTGAGGACGCCCCAGGGGACCTTGACCGTGGCCGCTGCTCGAGCGCTG 3894
Qy 4110 CTCAGGAGGAGGAGCTGGAGAGCCCCCAGTGCAGCATAGTCCCAAGGTTCTCGGCA 4169
Db 3895 CTCAGGAGGAGGAGCTGGAGAGCCCCCAGTGCAGCATAGTCCCAAGGTTCTCGGCA 3954

Qy 4170 CCAGAAGTTCTGCTGTAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4229
Db 3955 CCAGAAGTTCTGCTGTAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4014
Qy 4230 CCATCCCAGAGCTGCCAGTGTAGCCAGCCCGGTGGCCCGGCGCTGCTGCCAGTCTCCCG 4289
Db 4015 CCATCCCAGAGCTGCCAGTGTAGCCCGGCGCGGTGGCCCGGCGCTGCTGCCAGTCTCCCG 4074
Qy 4290 GCTGCAGCTGTGTGCTCCCTCCGCCAGCAGTGTGACCGGCTCTGGGGAGTGTGTCAG 4349
Db 4075 GCTGCAGCTGTGTGCTCCCTCCGCCAGCAGTGTGACCGGCTCTGGGGAGTGTGTCAG 4134
Qy 4350 AACCTGACAGGCCGGGAACCTCTCTGCTTCTGCTTCAAGACTACCCCGCTGCTGTGGCC 4409
Db 4135 AACCTGACAGGCCGGGAACCTCTCTGCTTCTGCTTCAAGACTACCCCGCTGCTGTGGCC 4194
Qy 4410 CAGGCTCTGAAGACTAAGAACTGAGTGTGATGAGGTGAGTACGAGGCTTCTGCTGGG 4469
Db 4195 CAGGCTCTGAAGACTAAGAACTGAGTGTGATGAGGTGAGTACGAGGCTTCTGCTGGG 4254
Qy 4470 GGCAGAGACCCAGAGCTGCCCTCCGGCCAAAGTGTGGCCGCTCAGTGCAGAGTGTGG 4529
Db 4255 GGCAGAGACCCAGAGCTGCCCTCCGGCCAAAGTGTGGCCGCTCAGTGCAGAGTGTGG 4314
Qy 4530 GCGCTGCTGAGTCCCTGCTGCGGGGCGCTCCGACCGTGTCTTGAAAAAACCCTCACAGCC 4589
Db 4315 GCGCTGCTGAGTCCCTGCTGCGGGGCGCTCCGACCGTGTCTTGAAAAAACCCTCACAGCC 4374
Qy 4590 TGGGCTCACAGCTGGATGTCTCAGGACAGTCTCAAGATCTGGTTCACAAACAAAGGCTGG 4649
Db 4375 TGGGCTCACAGCTGGATGTCTCAGGACAGTCTCAAGATCTGGTTCACAAACAAAGGCTGG 4434
Qy 4650 CACTCCATGGTGGGCTTTGTCAAACGAGCCAGCAACGCAATCTCTCGTGTCTCACTGCCC 4709
Db 4435 CACTCCATGGTGGGCTTTGTCAAACGAGCCAGCAACGCAATCTCTCGTGTCTCACTGCCC 4494
Qy 4710 CCAGGCCCGGCCCGCCAGCCACAGCATCACACACTCAACACCCCTTGAACCTCACC 4769
Db 4495 CCAGGCCCGGCCCGCCAGCCACAGCATCACACACTCAACACCCCTTGAACCTCACC 4554
Qy 4770 AAGAGACAGCTGTCTGAGGCTGCACTGATGCGCTCTCTCGTGGAGCTCTCGTCTCCATC 4829
Db 4555 AAGAGACAGCTGTCTGAGGCTGCACTGATGCGCTCTCTCGTGGAGCTCTCGTCTCCATC 4614
Qy 4830 TGTGTGCTCTTTGCCATGCTCTTTGTCCGCGCAGCTTCACTCTGTCTCTCATTCAGGAG 4889
Db 4615 TGTGTGCTCTTTGCCATGCTCTTTGTCCGCGCAGCTTCACTCTGTCTCTCATTCAGGAG 4674
Qy 4890 CGAGTCAACCCAGGACCAAGCACCTGAGCTCATGCGGGGCTGTGCCCCACCTCTACTGG 4949
Db 4675 CGAGTCAACCCAGGACCAAGCACCTGAGCTCATGCGGGGCTGTGCCCCACCTCTACTGG 4734
Qy 4950 CTTGGCAACTTTCTCTGGGACATGTAACTACTGTTGGTGGCAGCATGATCGTGGTGTCTC 5009
Db 4735 CTTGGCAACTTTCTCTGGGACATGTAACTACTGTTGGTGGCAGCATGATCGTGGTGTCTC 4794
Qy 5010 ATCTTTCTGGGCTTCCAGCAGAGGGCATATGTGGCCCTGCAACCTGCTGCTCTCTG 5069
Db 4795 ATCTTTCTGGGCTTCCAGCAGAGGGCATATGTGGCCCTGCAACCTGCTGCTCTCTG 4854
Qy 5070 CTGTTGTCTACTACTGATGCTGCTGATCACCGCTCATGTATCCAGCCCTCTCTTCTTC 5129
Db 4855 CTGTTGTCTACTACTGATGCTGCTGATCACCGCTCATGTATCCAGCCCTCTCTTCTTC 4914
Qy 5130 TTCTCCGTGGCCAGCAGAGCTATGTGGTGTCTCACTGCAATAAACCTCTTTATTGGCATC 5189
Db 4915 TTCTCCGTGGCCAGCAGAGCTATGTGGTGTCTCACTGCAATAAACCTCTTTATTGGCATC 4974
Qy 5190 AATGGAAGCATGGGCACTTTGCTGCTGAGCTCTTCTGATCAGAGCTGCAGAGGTG 5249
Db 4975 AATGGAAGCATGGGCACTTTGCTGCTGAGCTCTTCTGATCAGAGCTGCAGAGGTG 5034

Qy	5250	AGCCGGATCTTTGAAACAAGGTCCTTCTTATCTTTCCCCCACTTCTGCTTTGGCGCGGGGGCTC	5309
Db	5035	AGCCGGATCTTTGAAACAAGGTCCTTCTTATCTTTCCCCCACTTCTGCTTTGGCGCGGGGGCTC	5094
Qy	5310	ATTGACATGGTCGGAAACAGGCCATCGCTGATGCTCTTGGCCATGCTGGAGACAGGCAG	5369
Db	5095	ATTGACATGGTCGGAAACAGGCCATCGCTGATGCTCTTGGCCATGCTGGAGACAGGCAG	5154
Qy	5370	TTCCAGTCAACCCCTCGCTGGAGGTCGTGGCAAGAACTCTTTGGCCATGCTGATACAG	5429
Db	5155	TTCCAGTCAACCCCTCGCTGGAGGTCGTGGCAAGAACTCTTTGGCCATGCTGATACAG	5214
Qy	5430	GGGCCCCCTTCTTCTTTCACACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAG	5489
Db	5215	GGGCCCCCTTCTTCTTTCACACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAG	5274
Qy	5490	CCGAGGCTGAGGTCCTTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCTGAACCG	5549
Db	5275	CCGAGGCTGAGGTCCTTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCTGAACCG	5334
Qy	5550	GAGCGGCTGGTCCAAAGAGCCACCCAGGGGATGCTGTTGCTGCTAGGAACTTGACCAAG	5609
Db	5335	GAGCGGCTGGTCCAAAGAGCCACCCAGGGGATGCTGTTGCTGCTAGGAACTTGACCAAG	5394
Qy	5610	GTATACCTGGGCAGAGGATGCCAGCTGTTTGACCGCTGTGCTCTGGGGAATTCCTCCCTGGT	5669
Db	5395	GTATACCTGGGCAGAGGATGCCAGCTGTTTGACCGCTGTGCTCTGGGGAATTCCTCCCTGGT	5454
Qy	5670	GAGTGTTTTGGGCTGCTGGGTGTGAATGAGCAGGGGAAGACGTCCACGTTTCGATGGTG	5729
Db	5455	GAGTGTTTTGGGCTGCTGGGTGTGAATGAGCAGGGGAAGACGTCCACGTTTCGATGGTG	5514
Qy	5730	ACGGGGACACATTTGGCCAGCAGAGGGCGAGGCTGCTGTCGAGGCCACACGCTGGCCCGG	5789
Db	5515	ACGGGGACACATTTGGCCAGCAGAGGGCGAGGCTGCTGTCGAGGCCACACGCTGGCCCGG	5574
Qy	5790	GAACCCAGTGTGCGCACCTCAGCATGGGATGCTGCCCTCAATCCGATGCCATCTTTGAG	5849
Db	5575	GAACCCAGTGTGCGCACCTCAGCATGGGATGCTGCCCTCAATCCGATGCCATCTTTGAG	5634
Qy	5850	CTGCTGAGGGGCCGAGACACCTGGAGCTGCTTGGCGGCTGTGGCGGCTGTCCGGAGGCC	5909
Db	5635	CTGCTGAGGGGCCGCGAGACACCTGGAGCTGCTTGGCGGCTGTGGCGGCTGTCCGGAGGCC	5694
Qy	5910	CAGTTCCCGACACCGCTGCTCAGGCCTTGGCGGCTCTGCGACTCTCATGGTACGACAG	5969
Db	5695	CAGTTCCCGACACCGCTGCTCAGGCCTTGGCGGCTCTGCGACTCTCATGGTACGACAG	5754
Qy	5970	CGGCTGCAAGCACCTACGCGGAGGAAACAACGCAAGCTGGCGACGCGCCCTGGCGCTG	6029
Db	5755	CGGCTGCAAGCACCTACGCGGAGGAAACAACGCAAGCTGGCGACGCGCCCTGGCGCTG	5814
Qy	6030	GTTGGGACCCAGCCGTGTGTTTCTGGAAGAGCCGACACAGGCATGGAACCCAGCGCG	6089
Db	5815	GTTGGGACCCAGCCGTGTGTTTCTGGAAGAGCCGACACAGGCATGGAACCCAGCGCG	5874
Qy	6090	CGGCGCTTCTTTGGAAACAGCTTTTGGCGGTCGTGGCGGAGGGCCGTCAGTGAATGCTC	6149
Db	5875	CGGCGCTTCTTTGGAAACAGCTTTTGGCGGTCGTGGCGGAGGGCCGTCAGTGAATGCTC	5934
Qy	6150	ACCTCCCATAGCATGGAGGATGTGAACGCTCTGCTCGCGCCCTAGCCATCATGGTGAAT	6209
Db	5935	ACCTCCCATAGCATGGAGGATGTGAACGCTCTGCTCGCGCTAGCCATCATGGTGAAT	5994
Qy	6210	GGGCGGTTCCGCTGCTGGGAGCCCGGAAACATCTCAAGGGCAGATTTGCGGCGGGTTCAC	6269
Db	5995	GGGCGGTTCCGCTGCTGGGAGCCCGGCAACATCTCAAGGGCAGATTTGCGGCGGGTTCAC	6054
Qy	6270	ACACTGACCTTGGGGTCCCGCGCAAGGTCCTCCAGCCGACGCGCTTCGTGGGGGCC	6329
Db	6055	ACACTGACCTTGGGGTCCCGCGCAAGGTCCTCCAGCCGACGCGCTTCGTGGGGGCC	6114
Qy	6330	GAGTTCCCTGGGTCGAGCTGCGCGAGGCACATGGAGGCGCCCTGCGCTTCCAGCTGCCG	6389

Db	6115	GAGTTCCTCGGTGGAGCTGCGCGAGGCATGGGAGGTCCCTGCGCTTCCAGCTGCCG	6174
Qy	6390	CCGGGAGGGCGCTGCGCCCTGGCGCGCTCTTTGGGAGAGCTGCGCGGTGCACGGCGCAGAG	6449
Db	6175	CCGGGAGGGCGCTGCGCCCTGGCGCGCTCTTTGGGAGAGCTGCGCGGTGCACGGCGCAGAG	6234
Qy	6450	CACGGCTGGAGGACTTTTCCGTGAGCCAGCATGCTGGAGGAGGTATTCTTTGTACTTC	6509
Db	6235	CACGGCTGGAGGACTTTTCCGTGAGCCAGCATGCTGGAGGAGGTATTCTTTGTACTTC	6294
Qy	6510	TCCAGGACCAGGGGAAGGACGAGGACACCGAAGACGACGAAGGAGGAGGTGGGAGTG	6569
Db	6295	TCCAAGGACCAGGGGAAGGACGAGGACACCGAAGACGACGAAGGAGGAGGTGGGAGTG	6354
Qy	6570	GACCCCGCCGCGCTGCGACACCCCAACGCGTCAGCCAGTTCCTCGATGACCCCTAGC	6629
Db	6355	GACCCCGCCGCGCTGCGACACCCCAACGCGTCAGCCAGTTCCTCGATGACCCCTAGC	6414
Qy	6630	ACTGCCGAGACTGTGCTC	6647
Db	6415	ACTGCCGAGACTGTGCTC	6432
RESULT 8			
ABX72257			
ID	ABX72257 standard; cDNA; 6327 BP.		
AC	ABX72257;		
XX			
DT	03-JUN-2003 (first entry)		
XX			
DE	Human NOVX polynucleotide #88.		
XX			
KW	Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;		
KW	hypertension; congenital heart defect; aortic stenosis; valve disease;		
KW	atrial septal defect; atrioventricular canal defect; ductus arteriosus;		
KW	pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;		
KW	tubercous sclerosis; scleroderma; atherosclerosis; infectious disease;		
KW	obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;		
KW	Parkinson's disease; immune disorder; haematopoietic disorder;		
KW	haemophilia; hypercoagulation; Crohn's disease; cancer.		
OS	Homo sapiens.		
XX			
PN	W0200281498-A2.		
XX			
PD	17-OCT-2002.		
XX			
PF	03-APR-2002; 2002WO-US010780.		
XX			
PR	03-APR-2001; 2001US-0281086P.		
PR	03-APR-2001; 2001US-0281136P.		
PR	05-APR-2001; 2001US-0281863P.		
PR	05-APR-2001; 2001US-0281906P.		
PR	06-APR-2001; 2001US-0282020P.		
PR	10-APR-2001; 2001US-0282930P.		
PR	10-APR-2001; 2001US-0282934P.		
PR	12-APR-2001; 2001US-0283512P.		
PR	13-APR-2001; 2001US-0283710P.		
PR	17-APR-2001; 2001US-0284234P.		
PR	19-APR-2001; 2001US-0285325P.		
PR	20-APR-2001; 2001US-0285381P.		
PR	20-APR-2001; 2001US-0285609P.		
PR	23-APR-2001; 2001US-0285748P.		
PR	23-APR-2001; 2001US-0285890P.		
PR	24-APR-2001; 2001US-0286068P.		
PR	25-APR-2001; 2001US-0286292P.		
PR	27-APR-2001; 2001US-0287213P.		
PR	02-MAY-2001; 2001US-0288257P.		
PR	29-MAY-2001; 2001US-0294164P.		
PR	30-MAY-2001; 2001US-0294484P.		
PR	18-JUN-2001; 2001US-0298952P.		

Db 3541 |||||CGCCACAGGAGACAGCGCTGAGAAACGGGAAACAGCTGGGTACGCCCCAGAGACTGAC||| 3600
Qy 3810 CAGGCTCTGGGCCAGAGCCCGTGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG 3869
Db 3601 CAGGCTCTGGGCCAGAGCCCGTGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG 3660
Qy 3870 CTCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCCGAGCCGCCGGCCGTGTC 3929
Db 3661 CTCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCCGAGCCGCCGGCCGTGTC 3720
Qy 3930 GCCAGATCGTGTGCTGCTGCTTCTTGTGGCCCTGGCCCTGCTGCTCAGGCTCATGCTG 3989
Db 3721 GCCAGATCGTGTGCTGCTGCTTCTTGTGGCCCTGGCCCTGCTGCTCAGGCTCATGCTG 3780
Qy 3990 CCTCTTTTGGGCACTACCCGCTCTGGGCTCAGTCCACCATGACGCTGCTCAGGCTG 4049
Db 3781 CCTCTTTTGGGCACTACCCGCTCTGGGCTCAGTCCACCATGACGCTGCTCAGGCTG 3840
Qy 4050 TCCTTTTCAAGTGAAGACGCCCCAGGGACCCCTGGACGCTGCTGCTCAGGCGCTG 4109
Db 3841 TCCTTTTCAAGTGAAGACGCCCCAGGGACCCCTGGACGCTGCTGCTCAGGCGCTG 3900
Qy 4110 CTGAGGAGGAGGACTGGAGAGCCGCCAGTGCAGCATAGCTCCACAGGTTCTCGGCA 4169
Db 3901 CTGCAAGGAGGAGGACTGGAGAGCCGCCAGTGCAGCATAGCTCCCAAGGTTCTCGGCA 3960
Qy 4170 CCAGAACTTCTGCTGAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4229
Db 3961 CCAGAACTTCTGCTGAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCAGAGTCT 4020
Qy 4230 CCATCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4289
Db 4021 CCATCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
Qy 4290 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4349
Db 4081 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4140
Qy 4350 AACCTGACAGCCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4409
Db 4141 AACCTGACAGCCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4200
Qy 4410 CAGGCTGGAAGACTGAAGTGGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 4469
Db 4201 CAGGCTGGAAGACTGAAGTGGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 4260
Qy 4470 GCGCGAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4529
Db 4261 GCGCGAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4320
Qy 4530 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4589
Db 4321 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4380
Qy 4590 TGGGCTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4649
Db 4381 TGGGCTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4440
Qy 4650 CACTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4709
Db 4441 CACTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
Qy 4710 CCAGGCCCGGCCCGCCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGC 4769
Db 4501 CCAAGGCCCGGCCCGCCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGC 4560
Qy 4770 AAGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4829
Db 4561 AAGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4620
Qy 4830 TGTGTGCTTTTGCCATGCTCTTTTGTCCCGCCAGCTTCACTCTTGTCTCTCAATTGAGAG 4889
|||||

Db 4621 TGTGTGCTTTTGCCATGCTCTTTGTCCCGCCAGCTTCACTCTTGTCTCTCATTTGAGGAG 4680
Qy 4890 CGAGTCACTCCAGAGCCCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4949
Db 4881 CGAGTCACTCCAGAGCCCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4740
Qy 4950 CTGCGCAACTTCTCTGCGACATGTTAACTACTTGTGTGCGCAGCATGATCGTGGTGTCTC 5009
Db 4741 CTGCGCAACTTCTCTGCGACATGTTAACTACTTGTGTGCGCAGCATGATCGTGGTGTCTC 4763
Qy 5010 ATCTTTTGTGGCTTCCAGCAGAGGCGCATATGTGGCCCTGCGCAACCTGCTGCTGCTGCTG 5069
Db 4764 ----- 4763
Qy 5070 CTGTTGCTACTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5129
Db 4764 ----- 4763
Qy 5130 TTCTCGGTGCCAGCAGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5189
Db 4764 ----- 4763
Qy 5190 AATGGAAGCATGGCCACCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5249
Db 4764 ----- 4763
Qy 5250 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5309
Db 4780 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4839
Qy 5310 ATTGACATGCTGCGAAACAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5369
Db 4840 ATTGACATGCTGCGAAACAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4899
Qy 5370 TTCCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5429
Db 4900 TTCCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4959
Qy 5430 GGGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5489
Db 4960 GGGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5019
Qy 5490 CCCAGGCTGAGGTCTCTGCCACTCTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5549
Db 5020 CCCAGGCTGAGGTCTCTGCCACTCTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5079
Qy 5550 GAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5609
Db 5080 GAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5139
Qy 5610 GTATACCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5669
Db 5140 GTATACCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5199
Qy 5670 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5729
Db 5200 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5259
Qy 5730 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5789
Db 5260 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5319
Qy 5790 GAACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5849
Db 5320 GAACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5379
Qy 5850 CTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5909
Db 5380 CTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5439
Qy 5910 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5969
Db 5440 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5499

Db 5491 |||||CTTCTTTCACTACTCTCTGAGCACCGAAGCCAACTCTCTCCACACCCCGAGGGTGAGG 5550
Qy 5502 TCTCTGCCACTCTCTGGGAGAGGACGAGATGTAGCCCGTGAACGGGAGCGGTGGTC 5561
Db 5551 TCTCTGCCACTCTCTGGGAGAGGACGAGATGTAGCCCGTGAACGGGAGCGGTGGTC 5610
Qy 5562 CAAGAGACCAACCGAGGGGATGTGTGTGTCTGAGGAACTTGACCAAGGTATACCGTGGG 5621
Db 5611 CAAGAGACCAACCGAGGGGATGTGTGTGTCTGAGGAACTTGACCAAGGTATACCGTGGG 5670
Qy 5622 CAGAGGATGCCAGCTGTGACCGCTGTGCTGGGATTCCTCGGTGAGTGTTTTGGG 5681
Db 5671 CAGAGGATGCCAGCTGTGACCGCTGTGCTGGGATTCCTCGGTGAGTGTTTTGGG 5730
Qy 5682 CTGCTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTTCGATGGTGACGGGGACACA 5741
Db 5731 CTGCTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTTCGATGGTGACGGGGACACA 5790
Qy 5742 TTGGCCAGCAGGGGCGAGGCTGTGTGTGCGAGGCCACACGCTGGCCCGGGAACCCAGTGCT 5801
Db 5791 TTGGCCAGCAGGGGCGAGGCTGTGTGTGCGAGGCCACACGCTGGCCCGGGAACCCAGTGCT 5850
Qy 5802 CGGACACTCAGCATGGGATATCGCCCTCAATCCGATGCATCTTTGAGCTGCTGACGGGC 5861
Db 5851 CGGACACTCAGCATGGGATATCGCCCTCAATCCGATGCATCTTTGAGCTGCTGACGGGC 5910
Qy 5862 CGCGAGCACCTGGAGCTCTTTCGCGCGCTGCGCGGTGTCGCGAGGCCCGAGGTTGCCAG 5921
Db 5911 CGCGAGCACCTGGAGCTCTTTCGCGCGCTGCGCGGTGTCGCGAGGCCCGAGGTTGCCAG 5970
Qy 5922 ACCGCTGGCTCAGGCTGGCGCGCTGTGGGACTCTCATGTTACGACAGCCGCGCTGCAGGC 5981
Db 5971 ACCGCTGGCTCAGGCTGGCGCGCTGTGGGACTCTCATGTTACGACAGCCGCGCTGCAGGC 6030
Qy 5982 ACCTACAGCAGGGGAAACAAAGCAGCTGGGAGCGCCCTGGCGTGGTGGGGACCCA 6041
Db 6031 ACCTACAGCAGGGGAAACAAAGCAGCTGGGAGCGCCCTGGCGTGGTGGGGACCCA 6090
Qy 6042 GCGGTGGTGTCTTGGAGCGCCGACACACGACATGGACCCCGCGCGCGCTTCTCTT 6101
Db 6091 GCGGTGGTGTCTTGGAGCGCCGACACACGCGCATGGACCCCGCGCGCGCTTCTCTT 6150
Qy 6102 TGGAAACGCTTTTGGCGCTGGTGGGAGGCGCGTTCAGTCACTGCTCACCTCCCATAGC 6161
Db 6151 TGGAAACGCTTTTGGCGCTGGTGGGAGGCGCGTTCAGTCACTGCTCACCTCCCATAGC 6210
Qy 6162 ATGAGGAGTGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGCGGTTCCGC 6221
Db 6211 ATGAGGAGTGTGAAGCGCTCTGCTCGCGCTTAGCCATCATGTTGAATGGCGGTTCCGC 6270
Qy 6222 TGCTTGGGCGCCGACACATCTCAAGGCGAGATTCGCGCGGTTCAACACTGACCTG 6281
Db 6271 TGCTTGGGCGCCGACACATCTCAAGGCGAGATTCGCGCGGTTCAACACTGACCTG 6330
Qy 6282 CGGCTGCCCGCGCAAGTCCAGCCGCGAGCGGCTTCGTTGGGCGCGAGTTCCTCGG 6341
Db 6331 CGGCTGCCCGCGCAAGTCCAGCCGCGAGCGGCTTCGTTGGGCGCGAGTTCCTCGG 6390
Qy 6342 TCGAGTGTCCGAGGACATGAGAGCGCTGTCGCTTCCAGCTTCCGCGCGGAGGCGC 6401
Db 6391 TCGAGTGTCCGAGGACATGAGAGCGCTGTCGCTTCCAGCTTCCGCGCGGAGGCGC 6450
Qy 6402 TGCGCCCTGGCGCGCTTTTGGAGAGCTGCGGTGACGCGCGAGACACGCGGTGGAG 6461
Db 6451 TGCGCCCTGGCGCGCTTTTGGAGAGCTGCGGTGACGCGCGAGACACGCGGTGGAG 6510
Qy 6462 GACTTTTCCGTTGAGCAGACGATGCTGGAGAGGTATTCTTGTATCTTCCAAAGACCCAG 6521
Db 6511 GACTTTTCCGTTGAGCAGACGATGCTGGAGAGGTATTCTTGTATCTTCCAAAGACCCAG 6570
Qy 6522 GGGAGGACGAGGACACCGAAGAGCAGAGGAGCGAGGTGGAGTGGACCCCGCGCCA 6581
|||||

Db 6571 GGGAGGACGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGCGCA 6630
Qy 6582 GGCCTGAGCAGACCCCAACCGCTCAGCAGTTCTCTCGATGACCTTAGCACTGCCGAGACT 6641
Db 6631 GGCCTGAGCAGACCCCAACCGCTCAGCAGTTCTCTCGATGACCTTAGCACTGCCGAGACT 6690
Qy 6642 GTGCTC 6647
Db 6691 GTGCTC 6696
RESULT 10
ABX5284
ID ABX5284 standard; cDNA; 6651 BP.
XX
AC ABX5284;
XX
DT 18-JUN-2003 (first entry)
XX
cDNA encoding human ABCA7 splice variant #2.
XX
Human; ATP-binding cassette transporter protein A7; ABC transporter;
XX ABCA7; autoimmune disease; Sjogren's syndrome; inflammation;
XX abnormal lipid metabolism; arteriosclerosis; ABCA-SSN inhibitor;
XX immunomodulator; immunosuppressive; antiinflammatory;
XX antiarteriosclerotic; gene; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 673..6651
FT /*tag= a
FT /partial
FT /product= "ABCA7 splice variant #2"
FT /note= "This sequence lacks a stop codon. The ORF given
as SEQ ID No:5 is specifically claimed in Claim 4"
XX
PN WO2003010315-A1.
XX
PD 06-FEB-2003.
XX
PF 24-JUL-2002; 2002WO-JP007487.
XX
PR 25-JUL-2001; 2001JP-00224176.
PR 06-DEC-2001; 2001JP-00372530.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX
PI Ueda K, Nakagawa S, Nagase T;
XX
XX WPI; 2003-239444/23.
DR P-PSDB; ABU08465.
XX
PT Novel ABC transporter protein, ABCA7 splicing variant, participating in
the immune system, applicable in diagnosis of and screening drugs for
e.g. autoimmune diseases, Sjogren's syndrome and inflammations.
XX
PS Claim 4; Page 116-130; 183pp; Japanese.
XX
XX The present invention relates to the isolation of human ATP-binding
cassette (ABC) transporter protein A7 (ABCA7) splice variants, and the
XX polynucleotide sequences encoding them. The protein is applicable in the
XX diagnosis and screening of drugs for autoimmune diseases, Sjogren's
XX syndrome, inflammations, abnormal lipid metabolism and arteriosclerosis.
XX It may also be used in a method for screening ABCA-SSN inhibitors. The
XX present sequence encodes human ABCA7 splice variant #2
XX
SQ Sequence 6651 BP; 1096 A; 2155 C; 2138 G; 1262 T; 0 U; 0 Other;
Query Match 85.7%; Score 5833.2; DB 8; Length 6651;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 5893; Conservative 0; Mismatches 8; Indels 45; Gaps 1;

Qy	5082	CTGTATGGCTGGTGCATCATACACCGCTCATGTATCCCGAGCCTCTTCTTCTTCTCCGTGCC	5141
Db	5131	CTGTATGGCTGGTGCATCATACACCGCTCATGTATCCCGAGCCTCTTCTTCTTCTCCGTGCC	5190
Qy	5142	AGCACAGCCTATGTGGTGCTCACTGGCATAAA	5201
Db	5191	AGCACAGCCTATGTGGTGCTCACTGGCATAAA	5250
Qy	5202	GCCACCTTTGTGCTTGAGCTCTTCTCTGTATCAGAACTGCAGAGGTGAGCCGGATCTTG	5261
Db	5251	GCCACCTTTGTGCTTGAGCTCTTCTCTGTATCAGAACTGCAGAGGTGAGCCGGATCTTG	5310
Qy	5262	AAA	5321
Db	5311	AAA	5370
Qy	5322	CGGAACAGAGCCATGGCTGTATGCTTTGAGCGCTTGGGACAGAGGAGTTTCCAGTCA	5381
Db	5371	CGGAACAGAGCCATGGCTGTATGCTTTGAGCGCTTGGGACAGAGGAGTTTCCAGTCA	5430
Qy	5382	CTGGCTGGGAGGTGGTTCGGCAAGAACTCTTGGCCATGCTGATACAGGGGCCCTCTTC	5441
Db	5431	CTGGCTGGGAGGTGGTTCGGCAAGAACTCTTGGCCATGCTGATACAGGGGCCCTCTTC	5490
Qy	5442	CTTCTCTTACACTGCTGAGCACCGAAGCCACTCTCTGCCACAGCCACGGGTGAGG	5501
Db	5491	CTTCTCTTACACTGCTGAGCACCGAAGCCACTCTCTGCCACAGCCACGGGTGAGG	5550
Qy	5502	TCTCTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAA	5561
Db	5551	TCTCTGCCACTCTCTGGGAGAGGAGACGAGGATGTAGCCCGTGAA	5610
Qy	5562	CAAGAGCCACCCAGGGGATGTGTTGGTCTGAGGAACTTGTACCAAGGTATACCGTGGG	5621
Db	5611	CAAGAGCCACCCAGGGGATGTGTTGGTCTGAGGAACTTGTACCAAGGTATACCGTGGG	5670
Qy	5622	CAGAGGATGCCAGCTGTGTGACCGCTTGTGCTGGGGATTC	5681
Db	5671	CAGAGGATGCCAGCTGTGTGACCGCTTGTGCTGGGGATTC	5730
Qy	5682	CTGCTGGGTGTGAATGGAGCAGGAGAACGCTTCCAGTTCCGATGTGTGACGCGGGGACACA	5741
Db	5731	CTGCTGGGTGTGAATGGAGCAGGAGAACGCTTCCAGTTCCGATGTGTGACGCGGGGACACA	5790
Qy	5742	TTGGCCACGAGGGGCGAGGTGTGCTGGCAGGCCACAGCTGTGCGCCCGGGAAACCCAGTGCT	5801
Db	5791	TTGGCCACGAGGGGCGAGGTGTGCTGGCAGGCCACAGCTGTGCGCCCGGGAAACCCAGTGCT	5850
Qy	5802	CGCGACTCTCAGCATGGGATATCGCCCTCAATTCGATGCCACTTTTGAAGCTGCTGACGGGC	5861
Db	5851	CGCGACTCTCAGCATGGGATATCGCCCTCAATTCGATGCCACTTTTGAAGCTGCTGACGGGC	5910
Qy	5862	CGCGAGCATCTGGAGCTGCTTGGCGGCTGCGCGGTGTCCCGGAGGCCACAGGTGTC	5921
Db	5911	CGCGAGCATCTGGAGCTGCTTGGCGGCTGCGCGGTGTCCCGGAGGCCACAGGTGTC	5970
Qy	5922	ACCGCTGGCTCAGGCTGTGGCGCTGTGGGACTCTCATGGTACGACACCGGCTGTGAGGC	5981
Db	5971	ACCGCTGGCTCAGGCTGTGGCGCTGTGGGACTCTCTCATGGTACGACACCGGCTGTGAGGC	6030
Qy	5982	ACCTACAGCGAGGGAAACAAACGCAAGCTGGCGACCGCCCTGGCGCTGGTTGGGACCCA	6041
Db	6031	ACCTACAGCGAGGGAAACAAACGCAAGCTGGCGACCGCCCTGGCGCTGGTTGGGACCCA	6090
Qy	6042	GCCGTGGTGTCTTGGACGAGCCGACCAAGGATGACACCGACCGCGCGGCTGCTT	6101
Db	6091	GCCGTGGTGTCTTGGACGAGCCGACCAAGGATGACACCGACCGCGCGGCTGCTT	6150
Qy	6102	TGGAAACAGCCCTTTTGGCGGTGGTTCGGGAGGGCGGTTTCACTGATGCTCACTCCCATAGC	6161
Db	6151	TGGAAACAGCCCTTTTGGCGGTGGTTCGGGAGGGCGGTTTCACTGATGCTCACTCCCATAGC	6210

Qy	6162	ATGAGAGAGGTGAAAGCGCTCTGCTCGCGCCTTAGCCATCATGTGTGAATGGCGCGGTTC	CGC	6222			
Db	6211	ATGAGAGAGGTGAAAGCGCTCTGCTCGCGCCTAGCCATCATGTGTGAATGGCGCGGTTC	CGC	6270			
Qy	6222	TGCCTGGGAGCCGGCAACATCTCAAGGGCAGATTGCGCGCGGTACACACTGACCCCTG		6281			
Db	6271	TGCCTGGGAGCCGGCAACATCTCAAGGGC-----		6300			
Qy	6282	CGGCTGCCCGCGCAAGGTCCTCCAGCCGCGCAGCGGCTTTCGTGGCGGCGGAGTTCCCTGGG		6341			
Db	6301	-----AGGTCCTCCAGCGGAGCGGCTTTCGTGGCGGCGGAGTTCCCTGGG		6345			
Qy	6342	TCGGAGCTGCGCGAGGGCACATGGAGGCCGCTTCGCTTCAGCTGCGCGCGGAGGCGC		6401			
Db	6346	TCGGAGCTGCGCGAGGGCACATGGAGGCCGCTTCGCTTCAGCTGCGCGCGGAGGCGC		6405			
Qy	6402	TGGCCCTTGGCGCGCGCTCTTTGGAGAGCTGGCGGCTGCACGGCCACAGCACGCGCTGGAG		6461			
Db	6406	TGGCCCTTGGCGCGCGCTCTTTGGAGAGCTGGCGGCTGCACGGCCACAGCACGCGCTGGAG		6465			
Qy	6462	GACTTTTTCGTGAGCCAGACGATGCTGGAGAGGTATCTTTGTACTTCTCCAAAGGACCCAG		6521			
Db	6466	GACTTTTTCGTGAGCCAGACGATGCTGGAGAGGTATCTTTGTACTTCTCCAAAGGACCCAG		6525			
Qy	6522	GGGAAGGACGAGGACACCGAAGAGCAGAAAGGAGGAGGAGTGGGAGTGGACCCCGGCCCA		6581			
Db	6526	GGGAAGGACGAGGACACCGAAGAGCAGAAAGGAGGAGGAGTGGGAGTGGACCCCGGCCCA		6585			
Qy	6582	GGCCTGAGCACCCTCAAAACCGCTAGCCAGTTCTCTGATGACCTAGCACTGCGAGACT		6641			
Db	6586	GGCCTGAGCACCCTCAAAACCGCTAGCCAGTTCTCTGATGACCTAGCACTGCGAGACT		6645			
Qy	6642	GTGCTC	6647				
Db	6646	GTGCTC	6651				
RESULT 11							
AAD05626/c							
ID	AAD05626 standard; cDNA; 5811 BP.						
XX	AC						
XX	AC						
XX	XX						
DT	17-JUL-2001 (first entry)						
XX							
DE	Human secreted protein-encoding gene 11 cDNA clone HWBPE57, SEQ ID NO:58.						
XX							
KW	Human; secreted protein; proliferative disorder; cancer; tumour;						
KW	foetal abnormality; developmental abnormality; haematopoietic disorder;						
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;						
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;						
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;						
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;						
KW	cardiovascular disorder; angiogenic disorder; kidney disorder;						
KW	gastrointestinal disorder; pregnancy-related disorder;						
KW	endocrine disorder; infection; wound healing; vulvuary; cell culture;						
KW	chemotaxis; food additive; gene therapy; binding partner identification;						
SS							
XX							
OS	Homo sapiens.						
XX							
FH	Key	Location/Qualifiers					
FT	CDS	3300..3413					
FT		/*tag= a					
FT		/product= "Human secreted protein"					
FT		/note= "CDS does not include start codon"					
FT		/partial					
FT	sig_peptide	3300..3347					
FT		/*tag= b					
FT	mat_peptide	3348..3410					
FT		/*tag= c					
FT		/product= "Human mature secreted protein"					
XX	FT						

PN WO200134627-A1.
XX 17-MAY-2001.
XX 08-NOV-2000; 2000WO-US030628.
XX 12-NOV-1999; 99US-0164744P.
PR 30-JUN-2000; 2000US-0215140P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatoulis GA, Baker KP, Young PE,
XX WPI; 2001-316491/33.
DR P-PSDB; AAE01817.
XX New nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX Parkinson's diseases and cancers.
XX Claim 1; Page 457-458; 567pp; English.
XX AAD05579-AAD05658 represent cDNAs corresponding to 28 human secreted
XX protein genes and AAE01770-AAE01849 represent the proteins they encode.
XX AAE01850-AAE01860 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 28 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angionenic disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues, to identify their cognate ligands or binding
XX partners, and in chemotaxis, and can be used as a food additive or
XX preservative to modify storage properties. Antibodies specific for a
XX protein of the invention can be used in alleviating symptoms associated
XX with the disorders mentioned above, and in diagnostic immunoassays e.g.
XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX present sequence represents a human secreted protein-encoding cDNA of the
XX invention
XX
XX Sequence 5811 BP; 1068 A; 1890 C; 1831 G; 1006 T; 0 U; 16 Other;
Query Match 81.3%; Score 5529; DB 4; Length 5811;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 5709; Conservative 12; Mismatches 40; Indels 125; Gaps 4;
Qy 921 CTCACCTGGTACGAGGCTAGTACCTGATGAGCTGTGGGGCAGAGCCAGAAATCCGCC 980
Db |||||
Qy 5801 CTTCACTGGTACGANGCTAGTGACCTGATGAGCTGTGGGGCAGAGCCAGAAATCCGCC 5742
Db |||||
Qy 981 CTGCGACAGACAGCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 1040
Db |||||
Qy 5741 CTGCGAGACAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 5682
Db |||||
Qy 1041 CCGCTGTCCCGCTGCTCTGGAGACGCTCTGAAGCCTCTGTATCCTCGGGAAGCTACTCTTT 1100
Db |||||
Qy 5681 CCGCTGTCCCGCTGCTCTGGAGACGCTCTGAAGCCTCTGTATCCTCGGGAAGCTACTCTTT 5622
Db |||||
Qy 1101 GCACAGATACACTTTTATCCCGGAAGCTATGCGCCAGAGTGAACCGGACCTTCGAGGAG 1160
Db GCACAGATACACTTTTATCCCGGAAGCTATGCGCCAGAGTGAACCGGACCTTCGAGGAG 5562

Qy 1161 CTCACCTGTGCTGAGGAGTGTCCGGAGGTGTGGGAGATGTGGGACCCCGGATCTTTCACC 1220
Db |||||
Qy 5561 CTCACCTGTGCTGAGGAGTGTCCGGAGGTGTGGGAGATGTGGGACCCCGGATCTTTCACC 5502
Db |||||
Qy 1221 TTCAATGAACGACAGTTCATATGTGGCCATGTCTGACGCGGCTCTCTGACATGACAGATGAA 1280
Db |||||
Qy 5501 TTYATGAACGACAGTTCATATGTGGCCATGTCTGACGCGGCTCTCTGACATGACAGATGAA 5442
Db |||||
Qy 1281 GGAAGAAGCGAGCCAGACCTGGAGGCGGCGGACACACATGGAGGCGCTCCGATCTCTTCTG 1340
Db |||||
Qy 5441 GGAAGAAGCGAGCCAGACCTGGAGGCGGCGGACACACATGGAGGCGCTCTCGATCTCTTCTG 5382
Db |||||
Qy 1341 GACCTCTGGAGCGGTGGCTACAGCTGGCAGACGACACACGCTGATGTGGGGCACCTGGTG 1400
Db |||||
Qy 5381 GACCTCTGGAGCGGTGGCTACAGCTGGCAGACGACACACGCTGATGTGGGGCACCTGGTG 5322
Db |||||
Qy 1401 GGCACGCTGGGCGGAGTGACGAGTGTCTGCTTGGACAAGCTGGAGGCGGCGACCTCA 1460
Db |||||
Qy 5321 GGCACGCTGGGCGGAGTGACGAGTGTCTGCTTGGACAAGCTGGAGGCGGCGACCTCA 5262
Db |||||
Qy 1461 GAGCAGCGCTGGGTGCTCGCGGCGCTGCAACTGCTCGCGGAACATCGATCTTGGGCGCGC 1520
Db |||||
Qy 5261 GAGCAGCGCTGGGTGCTCGCGGCGCTGCAACTGCTCGCGGAACATCGATCTTGGGCGCGC 5202
Db |||||
Qy 1521 GTCTGCTTCTTGGGACCTGAGGACTCTTTCAGACCCCAACAGACGACCCCAACCCAGACCTG 1580
Db |||||
Qy 5201 GTCTGCTTCTTGGGACCTGAGGACTCTTTCAGACCCCAACAGACGACCCCAACCCAGACCTG 5142
Db |||||
Qy 1581 GGCCCGGCGCACGTCGCGCATCAAAATCGCATGACATTTGACGTGGTTCACGAGACCAAT 1640
Db |||||
Qy 5141 GGCCCGGCGCACGTCGCGCATCAAAATCGCATGACATTTGACGTGGTTCACGAGGACCAAT 5083
Db |||||
Qy 1641 AAGATCAGGAGACAGTGTGGGACCTTGGGACCTTGGGACCTTGGGACCTTGGGACCTGCGC 1700
Db |||||
Qy 5082 AAGATCAGGAGACAGTGTGGGACCTTGGGACCTTGGGACCTTGGGACCTTGGGACCTGCGC 5023
Db |||||
Qy 1701 TACGTGTGGGCGCGCTTCTGTGTACTCTGCAAGACCTGTGTGGAGCGTGCAGCGCTTCCGCTG 1760
Db |||||
Qy 5022 TACGTGTGGGCGCGCTTCTGTGTACTCTGCAAGACCTGTGTGGAGCGTGCAGCGCTTCCGCTG 4963
Db |||||
Qy 1761 CTCAGCGCGCCAAACCCCGGCGCGCTTCTGTGTACTCTGCAAGACGATGCCCTATCCGCTGAT 1820
Db |||||
Qy 4962 CTCAGCGCGCCAAACCCCGGCGCGCTTCTGTGTACTCTGCAAGACGATGCCCTATCCGCTGAT 4903
Db |||||
Qy 1821 GTGACGACGCTGTCTCTGCGTGTCTGACGCGGCTGCTGCGCTTCTCTCTGACGCTGCGC 1880
Db |||||
Qy 4902 GTGACGACGCTGTCTCTGCGTGTCTGACGCGGCTGCTGCGCTTCTCTCTGACGCTGCGC 4843
Db |||||
Qy 1881 TGGATCTACTCCGTGACACTGACAGTGAAGCCCGTGTGCGGAGAGAGAGAGAGAGAGAGAG 1940
Db |||||
Qy 4842 TGGATCTACTCCGTGACACTGACAGTGAAGCCCGTGTGCGGAGAGAGAGAGAGAGAGAGAG 4783
Db |||||
Qy 1941 CGGACACCATGTCGCGCATGGGGCTCAGCGCGCGGCTGTCTGCTAGGCTGCTGCTC 2000
Db |||||
Qy 4782 CGGACACCATGTCGCGCATGGGGCTCAGCGCGCGGCTGTCTGCTAGGCTGCTGCTC 4723
Db |||||
Qy 2001 AGCTGCTCGGGCCCTTCTGCTCAGCGCGCGGCTGTCTGCTGCTGCTCAAGCTGGG 2060
Db |||||
Qy 4722 AGCTGCTCGGGCCCTTCTGCTCAGCGCGCGGCTGTCTGCTGCTGCTCAAGCTGGG 4663
Db |||||
Qy 2061 GACATCTCTCCCTACAGCACCCCGGGGCTGTCTCTCTGCTTGTGGAGCGCTTCCGCGTG 2120
Db |||||
Qy 4662 GACATCTCTCCCTACAGCACCCCGGGGCTGTCTCTCTGCTTGTGGAGCGCTTCCGCGTG 4603
Db |||||
Qy 2121 GCCACGCTGACCCAGAGCTCTCTGCTCAGCGCGCTTCTTCTCCCGCGCAACCTGGCTGCG 2180
Db |||||
Qy 4602 GCCACGCTGACCCAGAGCTCTCTGCTCAGCGCGCTTCTTCTCCCGCGCAACCTGGCTGCG 4543
Db |||||
Qy 2181 GCCTGCGGCGGCTGCGCTTCTCTCTTCTACCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 2240
Db |||||
Qy 4542 GCCTGCGGCGGCTGCGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 4483

QY 2241 CGGACCGGTGCGCGGGTGGCCGCTGGCGCGAGCCTGCTGTGCGCCGTGGCC-TT 2299
Db 4482 CGGGACCGGTGCGCGGGTGGCGGGTGGCGCGAGCCTGCTGTGCGCCGTGGCCTTT 4423
QY 2300 CGGCTTGGCTGGCG-AGAGCCTGGCTGTGTGAGGAGCAGGCGAGGGCGCGAGTGGC 2358
Db 4422 CGGCTTGGCTGGCGAGAGCCTGGCTGTGTGAGGAGCAGGCGAGGGCGCGAGTGGC 4363
QY 2359 ACAAAGTGGGACCCGGCCTACGGCAGACGCTCTTACGCTGGGCCAGGTCTTGGCCTTC 2418
Db 4362 ACAAAGTGGGACCCGGCCTACGGCAGACGCTCTTACGCTGGGCCAGGTCTTGGCCTTC 4303
QY 2419 TGTGCTGGAGCGGGCGCTCTAGCGCTCGCCACCTGTGTACCTGGAGCTGTGTGCCAG 2478
Db 4302 TGTGCTGGAGCGGGCGCTCTAGCGCTCGCCACCTGTGTACCTGGAGCTGTGTGCCAG 4243
QY 2479 GCAGTACGGGATCCCTGAAACCATGGAATTTCTTTTCGGAGGAGCTACTGTGGGAC 2538
Db 4242 GCCAGTACGGGATCCCTGAAACCATGGAATTTCTTTTCGGAGGAGCTACTGTGGGAC 4183
QY 2539 CTGGGCCCCCAAGATCGAGCCCTTGGCCACCCCGCTGGACCCAAAGGTGCTGGTAG 2598
Db 4182 CTGGGCCCCCAAGATCGAGCCCTTGGCCACCCCGCTGGACCCAAAGGTGCTGGTAG 4123
QY 2599 AAGAGCACCGCCCGGCTGAGTCTCGGCGTATCCGTTGCGAGCTGGAGAGCGCTTTC 2658
Db 4122 AAGAGCACCGCCCGGCTGAGTCTCGGCGTATCCGTTGCGAGCTGGAGAGCGCTTTC 4063
QY 2659 CTGGAAGCCCGACCGCCCTCGGGGGCTCAGCTGGACTTCTACAGGGCCACATCA 2718
Db 4062 CTGGAAGCCCGACCGCCCTCGGGGGCTCAGCTGGACTTCTACAGGGCCACATCA 4003
QY 2719 CGGCTTCTGGGGCCAAACGGGGCGGCAAGACCAACCCCTGTCCATCTTGAAGTGCC 2778
Db 4002 CGGCTTCTGGGGCCAAACGGGGCGGCAAGACCAACCCCTGTCCATCTTGAAGTGCC 3943
QY 2779 TCTTCCACCAAGTGGTGTCTGCTTTCATCTGCGGCCACGAGCTCCGCTCCAGCATGG 2838
Db 3942 TCTTCCACCAAGTGGTGTCTGCTTTCATCTGCGGCCACGAGCTCCGCTCCAGCATGG 3883
QY 2839 CCGCCATCCGGCCCACTGGGCGTCTGCTCAGTACAAAGTGTGTTTGACATGCTGA 2898
Db 3882 CCGCCATCCGGCCCACTGGGCGTCTGCTCAGTACAAAGTGTGTTTGACATGCTGA 3823
QY 2899 CCGTGGACGAGACCGTCTGTGACGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGA 2958
Db 3822 CCGTGGACGAGACCGTCTGTGCTGTGTTCTATGGGCGGTGAAGGCTCTGAGTGGCCGTGTAGTGG 3763
QY 2959 GCCCGAGCAGGACCGTCTGTGACGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGA 3018
Db 3762 GCCCGAGCAGGACCGTCTGTGACGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGA 3703
QY 3019 CTCGCCACCTCTGTGTTGGATGCAACGGAAGTGTGCTGGGCCATTCCTTTTGGGGCG 3078
Db 3702 CTCGCCACCTCTGTGTTGGATGCAACGGAAGTGTGCTGGGCCATTCCTTTTGGGGCG 3643
QY 3079 GCTCCCAAGTGTGTATCTGGAAGAGCTACGGCTGCGGTGGATCTGCTTCCCGCGCG 3138
Db 3642 KRCGCCAAGTGTGTATCTGGAAGAGCTACGGCTGCGGTGGATCTGCTTCCCGCGCG 3583
QY 3139 GTATTTGGAGTGTGCTCAAAATACGAGAAGTGCACGCTGATCCTCTCCACCCACC 3198
Db 3582 GTATTTGGAGTGTGCTCAAAATACGAGAAGTGCACGCTGATCCTCTCCACCCACC 3523
QY 3199 ACCTGGATGAGCAGAGCTGCTGGGACCGTGTGGCTGTGGTGGCAGTGGCCGCTTGT 3258
Db 3522 ACCTGGATGAGCAGAGCTGCTGGGACCGTGTGGCTGTGGTGGCAGTGGCCGCTTGT 3463
QY 3259 GCTGCTGTGCTCCCACTTCTGCGCGCTCACCTGGGCTCGGCTACTACTGACGC 3318
Db 3462 GCTGCTGTGCTCCCACTTCTGCGCGCTCACCTGGGCTCGGCTACTACTGACGC 3403
QY 3319 TGGTGAAGGCGCGCTGCCCCGTGACCAATGAGAAGGCTGACATGAGGAGGCA 3378

Db 3402 TGGTGAAGGCGCGCTGCCCCGTGACCAATGAGAAGGCTGACATGAGAGGCA 3343
QY 3379 GTGTGACACACAGGCAAGAAAAGAAATGGCAGCAGGGCAGCAGTGGCAGTCTCTC 3438
Db 3342 GTGTGACACACAGGCAAGAAAAGAAATGGCAGCAGGGCAGCAGTGGCAGTCTCTC 3283
QY 3439 AGCTGCTGGCCCTGTGTACAGCAGTGGTGCCTGGGACAGGCTGTGGAGAGCTGCCAC 3498
Db 3282 AGCTGCTGGCCCTGTGTACAGCAGTGGTGCCTGGGACAGGCTGTGGAGAGCTGCCAC 3223
QY 3499 ACAGCTGTGTGTGTGCTGCCCTACACGGGTGCCCATGACGGCAGTTCGCCACACTCT 3558
Db 3222 ACAGCTGTGTGTGTGCTGCCCTACACGGGTGCCCATGACGGCAGTTCGCCACACTCT 3163
QY 3559 TCCGAGAGCTAGACACCGGGCTGGCGAGCTGAGGTCTACTGGCTACGGGATCTCCGACA 3618
Db 3162 TCCGAGAGCTAGACACCGGGCTGGCGAGCTGAGGTCTACTGGCTACGGGATCTCCGACA 3103
QY 3619 CCAGCCTCGAGGAGATCTTCCCTGAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGG 3678
Db 3102 CCAGCCTCGAGGAGATCTTCCCTGAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGG 3043
QY 3679 AGGATGGCAGCTGCGGGCAGCACCTATGCACAGGCAATTCCTGGCTTAGACCTAACCTGC 3738
Db 3042 AGGATGGCAGCTGCGGGCAGCACCTATGCACAGGCAATTCCTGGCTTAGACCTAACCTGC 2983
QY 3739 GGCTCAAGATGCCGCCACAGGAGACAGCGCTGGAGAAACGGGAAACAGCTGGGTGAGCCC 3798
Db 2982 GGCTCAAGATGCCGCCACAGGAGACAGCGCTGGAGAAACGGGAAACAGCTGGGTGAGCCC 2923
QY 3799 CAGAGTACAGCAGGCTCTGGGCCAGACCGCTGGGCCGGGTACAGGGCTGGGCACTGA 3858
Db 2922 CAGAGTACAGCAGGCTCTGGGCCAGACCGCTGGGCCGGGTACAGGGCTGGGCACTGA 2863
QY 3859 CCGCCAGCAGCTTCCAGGGCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCGCAGCGGCC 3918
Db 2862 CCGCCAGCAGCTTCCAGGGCCTGCTTCTCAAGCGGNGGGGGGGGGGGGGGGG 2803
QY 3919 GGCGCCTGTTGGCCAGATGCTGCTGCCCTCTTTTGTGGCCCTGGGCCCTCGTGTCA 3978
Db 2802 GGCGCCTGCT----- 2793
QY 3979 GCCTCATGCTGCTCTTTTCGGGCACTACCGGCTCTGGGGCTCAGTCCACCATGTAAG 4038
Db 2792 ----- 2793
QY 4039 GTGCTCAGGTGCTTCTTTCAGTGAAGACGCCCCAGGGGACCTTGAGACGTGCCGGCTGC 4098
Db 2792 -----GTTCCGCCAGTGAAGACGCCCCAGGGGACCTTGAGACGTGCCGGCTGC 2745
QY 4099 TCAGGCGCTGTGTCAGGAGGACGAGCTGGAGGAGCCCCCAGTGCAGCATAGTCTCCACA 4158
Db 2744 TCAGGCGCTGTGTCAGGAGGACGAGCTGGAGGAGCCCCCAGTGCAGCATAGTCTCCACA 2685
QY 4159 GGTTCGCGCACAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGA 4218
Db 2684 GGTTCGCGCACAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGA 2625
QY 4219 CCCAGAGTCTCATCCCCCAGCCTGCTGAGTGCAGGAGCCCCGCTGCCCGGCTGTGTC 4278
Db 2624 CCCAGAGTCTCATCCCCCAGCCTGCTGAGTGCAGGAGCCCCGCTGCCCGGCTGTGTC 2565
QY 4279 CCGACTGCCCGGTGACGTGTGTGTCCTCCCTCGGCCAGGAGTGAACGGCTCTGGGG 4338
Db 2564 CCGACTGCCCGGTGACGTGTGTGTCCTCCCTCGGCCAGGAGTGAACGGCTCTGGGG 2505
QY 4339 AAGTGGTTCAGAACCTGACAGGCGGAACCTGTCTGACTTCTCTGCTCAAGACCTACCCGC 4398
Db 2504 AAGTGGTTCAGAACCTGACAGGCGGAACCTGTCTGACTTCTCTGCTCAAGACCTACCCGC 2445
QY 4399 GCCTGTGCGCCAGGCGCTGAAGACTTAAGAAGTGGGTGAATGAGGTACAGGTACGGAGCT 4458

QY 6619 ATGACCTAGCACTGCGAGACTGTGCTCTGAGCTCCCTCCCTGCGGGCGCGGGGA 6678
Db 224 ATGACCTAGCACTGCGAGACTGTGCTCTGAGCTCCCTCCCTGCGGGCGCGGGGA 165
QY 6679 GGCCTTGGGAATGCAAGGCAAGGTAGAGTGTGCTAGGAGCCCTGGAGCTGCGCAG 6738
Db 164 GGCCTTGGGAATGCAAGGCAAGGTAGAGTGTGCTAGGAGCCCTGGAGCTGCGCAG 105
QY 6739 AGGGCTGTGCTGCTGAGAAATTAAGAGAGGCTGGAGAGCGGTGCTGTGAAA 6798
Db 104 AGGGCTGTGCTGCTGAGAAATTAAGAGAGGCTGGAGAGCGGTGCTGTGAAA 45
QY 6799 AAAAAA 6804
Db 44 AAAAAA 39

RESULT 12

ADA40555/c

ID ADA40555 standard; cDNA; 5811 BP.

XX

AC ADA40555;

XX 20-NOV-2003 (first entry)

DT

XX Human secreted protein encoding cDNA.

DE

XX Human; secreted protein; cancer; hyperproliferative disorder;

KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;

KW anaemia; allergic reaction; asthma; cardiovascular disorder;

KW wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;

KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;

KW vulnerable; cardiac; gene therapy; ss.

XX Homo sapiens.

OS

XX WO2002102993-A2.

PN

XX 27-DEC-2002.

XX

XX 19-MAR-2002; 2002WO-US008123.

PF

XX 21-MAR-2001; 2001US-0277340P.

PR

XX 19-JUL-2001; 2001US-0306171P.

PR

XX 13-NOV-2001; 2001US-0331487P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX Rosen CA, Ruben SM;

XX WPI, 2003-175238/17.

DR

XX New human secreted proteins and nucleic acid molecules, useful for

PT preparing a diagnostic or pharmaceutical composition for diagnosing,

PT preventing or treating cancer or other hyperproliferative disorder,

PT asthma, allergies or AIDS.

XX

PS Claim 9; SEQ ID NO 937; 3205pp; English.

XX

XX The invention relates to novel genes ADA39629-ADA40565 and proteins

CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,

CC treating or ameliorating medical conditions e.g. by protein or gene

CC therapy. The polypeptides, nucleic acid molecules, antibodies or their

CC fragments, and agonists or antagonists that bind to the polypeptide are

CC useful for preparing a diagnostic or pharmaceutical composition for

CC diagnosing or treating cancer or other hyperproliferative disorder. The

CC polypeptides and nucleic acid molecules are also useful for detecting,

CC preventing, diagnosing, prognosticating, treating or ameliorating cancer

CC or other hyperproliferative disorders including neoplasms, autoimmune

CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus

CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic

CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,

CC thrombocytopenia), allergic reactions including asthma or eczema,

CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 5811 BP; 1068 A; 1890 C; 1831 G; 1006 T; 0 U; 16 Other;

Query Match

Best Local Similarity 81.3%; Score 5529; DB 8; Length 5811;

Matches 5709; Conservative 12; Mismatches 40; Indels 125; Gaps 4;

QY 921 CTCAACTGTAGAGGCTAGTGACCTGATGGAGCTGGTGGGGCAGAGCCAGAAATCGCC 980

Db 5801 CTTCACTGTAGGANGCTAGTGACCTGATGGAGCTGGTGGGGCAGAGCCAGAAATCGCC 5742

QY 981 CTGCCAGACAGAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 1040

Db 5741 CTGCCAGACAGAGCCTGAGCCCGCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCAC 5682

QY 1041 CGCTGTCCCGCTGCTCTGGAGACCCCTGAAGCCTCTGATCTCTGGGAAGCTACTCTTT 1100

Db 5681 CGCTGTCCCGCTGCTCTGGAGACCCCTGAAGCCTCTGATCTCTGGGAAGCTACTCTTT 5622

QY 1101 GCACAGATACACCTTTTACCCTGGAGAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAG 1160

Db 5621 GCACAGATACACCTTTTACCCTGGAGAGCTCATGGCCAGGTGAACCGGACCTTCGAGGAG 5562

QY 1161 CTCACCTGCTGAGGAGTCTCCGGAGGTGTCGGAGATCTGGGACCCCGGATCTTCACC 1220

Db 5561 CTCACCTGCTGAGGAGTCTCCGGAGGTGTCGGAGATCTGGGACCCCGGATCTTCACC 5502

QY 1221 TTCAATGAACGACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCTGCAGATGCAAGATGAA 1280

Db 5501 TTCAATGAACGACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCTGCAGATGCAAGATGAA 5442

QY 1281 GGAAGAAGCGAGCCGAGACCTGAGAGCCCGGACCAATGAGAGCCCTCGATCTCTCTG 1340

Db 5441 GGAAGAAGCGAGCCGAGACCTGAGAGCCCGGACCAATGAGAGCCCTCGATCTCTCTG 5382

QY 1341 GACCTTGGAGCGGTGGCTACAGCTGGCAGAGCGCACACGCTGATGTGGGACCTGGTG 1400

Db 5381 GACCTTGGAGCGGTGGCTACAGCTGGCAGAGCGCACACGCTGATGTGGGACCTGGTG 5322

QY 1401 GGCACGCTGGGCGAGTGACGAGTGCCTGTCTTGGCAAGCTGGAGCGGACCCCTCA 1460

Db 5321 GGCACGCTGGGCGAGTGACGAGTGCCTGTCTTGGCAAGCTGGAGCGGACCCCTCA 5262

QY 1461 GAGGACGCTGTGTGTCGGGCGCTTGAATCTGCTCGGGAACATCGATTTCTGGGCGGC 1520

Db 5261 GAGGACGCTGTGTGTCGGGCGCTTGAATCTGCTCGGGAACATCGATTTCTGGGCGGC 5202

QY 1521 GTGCTTTCTTGGGACCTGAGGACTTTTCAGACCCCAAGAGAGCCCAACCCAGACCTG 1580

Db 5201 GTGCTTTCTTGGGACCTGAGGACTTTTCAGACCCCAAGAGAGCCCAACCCAGACCTG 5142

QY 1581 GGGCCCGGCACTGTGGCATCAAAATCCGATGACATTTGACGTGTCTACAGGAGCAAT 1640

Db 5141 GGGCCCGG-CACTGTGGCATCAAAATCCGATGACATTTGACGTGTCTACAGGAGCAAT 5083

QY 1641 AAGATCAGGAGCAGGTTTGGGACCTTGCCCAAGCGCGGACCCCTGACCGGACCTGGCG 1700

Db 5082 AAGATCAGGAGCAGGTTTGGGACCTTGCCCAAGCGCGGACCCCTGACCGGACCTGGCG 5023

QY 1701 TACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCGTCCGCGTG 1760
DB 5022 TACGTGTGGGGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCGTCCGCGTG 4963
QY 1761 CTCACGGCGCCCAACCCCGGGCGGCTCTACCTGCGAGCAGATGCCCTATCCGTGCTAT 1820
DB 4962 CTCACGGCGCCCAACCCCGGGCGGCTCTACCTGCGAGCAGATGCCCTATCCGTGCTAT 4903
QY 1821 GTGGACGAGCTGTTCCTCGGTGTGTGAGCCGGTCCGTCCGCTCTTCTGAAGCTGGCC 1880
DB 4902 GTGGACGAGCTGTTCCTCGGTGTGTGAGCCGGTCCGTCCGCTCTTCTGAAGCTGGCC 4843
QY 1881 TGGATCTACTCCGTGACACTGACAGTGAAGCCGTGGTCCGGGAGAGAGAGCGCGCTG 1940
DB 4842 TGGATCTACTCCGTGACACTGACAGTGAAGCCGTGGTCCGGGAGAGAGAGCGCGCTG 4783
QY 1941 CGGGACACATGCGCGCCATGGGGCTCAGCGCGCGGTGCTGTGGCTAGGCTGGTTCCTC 2000
DB 4782 CGGGACACATGCGCGCCATGGGGCTCAGCGCGCGGTGCTGTGGCTAGGCTGGTTCCTC 4723
QY 2001 AGCTGCTCGGGCCCTTCTGTCTCAGCGCGCGGTGCTGTGGTCTCAAGCTGGGG 2060
DB 4722 AGCTGCTCGGGCCCTTCTGTCTCAGCGCGCGGTGCTGTGGTCTCAAGCTGGGG 4663
QY 2061 GACATCTCCCTACAGCAACCGGGCGTGTCTTCTGTCTGGCAGCGCTTCGCGTG 2120
DB 4662 GACATCTCCCTACAGCAACCGGGCGTGTCTTCTGTCTGGCAGCGCTTCGCGTG 4603
QY 2121 GCCAGGTGACCCAGAGCTTCTGTCTCAGCGCGCTTCTTCCCGCGCAACCTGGCTGCG 2180
DB 4602 GCCAGGTGACCCAGAGCTTCTGTCTCAGCGCGCTTCTTCCCGCGCAACCTGGCTGCG 4543
QY 2181 GCCTGCGCGCGCTGGCCCTACTTCTCCCTCTACCTGCGCTACGTGCTGTGTGGCTTGG 2240
DB 4542 GCCTGCGCGCGCTGGCCCTACTTCTCCCTCTACCTGCGCTACGTGCTGTGTGGCTTGG 4483
QY 2241 CGGACCGCGCTGCGCGGGTGGCGCGTGGCGCGAGCTGTGTGCCCTGCGC--TT 2299
DB 4482 CGGACCGCGCTGCGCGGGTGGCGCGTGGCGCGAGCTGTGTGCCCTGCGC--TT 4423
QY 2300 CGGCTTCGGCTGCG--AGAGCTTGGCTGTGTGGAGGAGCGAGGGCGCGAGTGGC 2358
DB 4422 CGGCTTCGGCTGCGAGAGCTTGGCTGTGTGGAGGAGCGAGGGCGCGAGTGGC 4363
QY 2359 ACAAGTGGGCAACCGGGCTACGGCAGACGCTTTCAGCTGCGCCAGCTCTGTGCGCTTC 2418
DB 4362 ACAAGTGGGCAACCGGGCTACGGCAGACGCTTTCAGCTGCGCCAGCTCTGTGCGCTTC 4303
QY 2419 TGCTGTGGACGCGCGCTCTACCGCTCGCAACCTGGTACCTGGAAGCTGTGTGCCAG 2478
DB 4302 TGCTGTGGACGCGCGCTCTACCGCTCGCAACCTGGTACCTGGAAGCTGTGTGCCAG 4243
QY 2479 GCCAGTACGGATCCCTGAACATGGAATTTTCTTTTCGAGAGCTACTGTGTGGAC 2538
DB 4242 GCCAGTACGGATCCCTGAACATGGAATTTTCTTTTCGAGAGCTACTGTGTGGAC 4183
QY 2539 CTCGGCCCCCAAGAGTCCAGCCCTTCCCGCCAGCTGCAACCAAGGTGTGTAG 2598
DB 4182 CTCGGCCCCCAAGAGTCCAGCCCTTCCCGCCAGCTGCAACCAAGGTGTGTAG 4123
QY 2599 AAGAGCACCGCGCGCTGAGTCTGCGGTATCCGTTTCGAGCTGAGAGAGCGCTTTC 2658
DB 4122 AAGAGCACCGCGCGCTGAGTCTGCGGTATCCGTTTCGAGCTGAGAGAGCGCTTTC 4063
QY 2659 CTGAAGCCCGACGACCTGTGGGGGCTGAGCTGGAATTTCTACAGGGCCACATCA 2718
DB 4062 CTGAAGCCCGACGACCTGTGGGGGCTGAGCTGGAATTTCTTACAGGGCCACATCA 4003
QY 2719 CCGCTTCTGGGCCACAAAGGGGCGGCAAGACCAACCTGTGCTATCTGATGGCC 2778
DB 4002 CCGCTTCTGGGCCACAAAGGGGCGGCAAGACCAACCTGTGCTATCTGATGGCC 3943
QY 2779 TCTTCCCAACCGAGTGTGGCTGTGCTTCTCATCTTGGGCCACAGAGCTCGCTCCAGCATGG 2838

DB 3942 TCTTCCCAACCGAGTGTGGCTGTGCTTTCATCTGGGCCACAGCTCGCTCCAGCATGG 3883
QY 2839 CCGCATCCGGCCCCCACCCTGGCGCTGTCTCTCAGTACAAACGTGTGTGTGACATGCTGA 2898
DB 3882 CCGCATCCGGCCCCCACCCTGGCGCTGTCTCTCAGTACAAACGTGTGTGTGACATGCTGA 3823
QY 2899 CCGTGGACGACACGCTGTGTGTCTATGGCGCGCTGAAAGGGTGTGAGTCCGCTGATGG 2958
DB 3822 CCGTGGACGACACGCTGTGTGTCTATGGCGCGCTGAAAGGGTGTGAGTCCGCTGATGG 3763
QY 2959 GCCCGAGACGACCGCTGTGTGAGGATGTGGGGCTGGTCTCCAGCAGAGTGTGACGA 3018
DB 3762 GCCCGAGACGACCGCTGTGTGAGGATGTGGGGCTGGTCTCCAGCAGAGTGTGACGA 3703
QY 3019 CTGCGCACTCTCTGTGTGGATGCAACGGAAGCTGTCCGTGGCCATTCCTCTTGTGGCG 3078
DB 3702 CTGCGCACTCTCTGTGTGGATGCAACGGAAGCTGTCCGTGGCCATTCCTCTTGTGGCG 3643
QY 3079 GCTCCCAAGTTGTTATTCCTGACGAGCTACCGGTGGCGTGGATCCTGCTTCCCGCCGCG 3138
DB 3642 KCKCCCAAGTTGTTATTCCTGACGAGCTACCGGTGGCGTGGATCCTGCTTCCCGCCGCG 3583
QY 3139 GTATTGGGAGCTGCTGCTCAATACCGAGAGGTCGACGCTGATCCTCTCCACCAC 3198
DB 3582 GTATTGGGAGCTGCTGCTCAATACCGAGAGGTCGACGCTGATCCTCTCCACCAC 3523
QY 3199 ACCTGGATGAGCGAGAGCTGTGGGAGACCGTGTGGCTGTGTGGAGTGGCGCTTGT 3258
DB 3522 ACCTGGATGAGCGAGAGCTGTGGGAGACCGTGTGGCGCTGTGTGGAGTGGCGCTTGT 3463
QY 3259 GCTGTGTGGCTCCCGCACTTCTTCTGCGCCCTACCTGGGCTTCCGGCTACTACCTGA 3318
DB 3462 GCTGTGTGGCTCCCGCACTTCTTCTGCGCGCTACCTGGGCTTCCGGCTACTACCTGA 3403
QY 3319 TGGTGAAGGCCCGCTGTGCTTGAACCAATGAGAGGCTGACACTGACATGAGAGGCA 3378
DB 3402 TGGTGAAGGCCCGCTGTGCTTGAACCAATGAGAGGCTGACACTGACATGAGAGGCA 3343
QY 3379 GTGTGGACACAGCGAGGAAAGAAAGTGCACCGCGGAGGAGAGTGGCGCTTCTC 3438
DB 3342 GTGTGGACACAGCGAGGAAAGAAAGTGCACCGCGGAGGAGAGTGGCGCTTCTC 3283
QY 3439 AGCTGTGGCGCTGTGCTACAGCACTGGGTGCGCGGGGACCGCTGGTGGAGAGCTGCCAC 3498
DB 3282 AGCTGTGGCGCTGTGCTACAGCACTGGGTGCGCGGGGACCGCTGGTGGAGAGCTGCCAC 3223
QY 3499 ACAGCTGTGTGTGCTGCTTACACCGGTGCCATGACGGGAGCTTTCGCCACACTCT 3558
DB 3222 ACAGCTGTGTGTGCTGCTTACACCGGTGCCATGACGGGAGCTTTCGCCACACTCT 3163
QY 3559 TCCGAGAGCTAGACACCGGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3618
DB 3162 TCCGAGAGCTAGACACCGGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3103
QY 3619 CCAGCTCGAGAGATCTTCTGAAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGG 3678
DB 3102 CCAGCTCGAGAGATCTTCTGAAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGG 3043
QY 3679 AGGATGGCAGCTGCGGGCAGCACCTATGCAAGGCAATTTGTGGCTAGACGTAACMTAC 3738
DB 3042 AGGATGGCAGCTGCGGGCAGCACCTATGCAAGGCAATTTGTGGCTAGACGTAACMTAC 2983
QY 3739 GGCTCAAGATGCCGCCACAGAGACAGCGCTGGAGAACCGGGGACACAGCTGGGTGAGGCC 3798
DB 2982 GGCTCAAGATGCCGCCACAGAGACAGCGCTGGAGAACCGGGGACACAGCTGGGTGAGGCC 2923
QY 3799 CAGAGATGACCGAGGCTCTGGGGCCAGACCGCTGGGGCCGGGTACAGGGCTGGGCACTGA 3858
DB 2922 CAGAGATGACCGAGGCTCTGGGGCCAGACCGCTGGGGCCGGGTACAGGGCTGGGCACTGA 2863
QY 3859 CCGGCCACAGAGCTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCGCAGCGGCC 3918

QY 3259 GCTGCTGTGGCTCCCACTCTTCTGCGCGCTCACTGGCTCCGGCTACTACTGACGC 3318
DB 3462 GCTGTGTGGTCCCACTCTTCTGCGCGCTCACTGGCTCCGGCTACTACTGACGC 3403
QY 3319 TGGTGAAGCGCGCTGCGCTGACCAACCAATGAGAGCTGACACTGACATGAGAGGCA 3378
DB 3402 TGGTGAAGCGCGCTGCGCTGACCAACCAATGAGAGCTGACACTGACATGAGAGGCA 3343
QY 3379 GTGTGGACACGAGGAGGAAAGAAATGGCAGCGAGGCAAGAGTGGCACTCTCTC 3438
DB 3342 GTGTGGACACGAGGAGGAAAGAAATGGCAGCGAGGCAAGAGTGGCACTCTCTC 3283
QY 3439 AGCTGCTGGCTGTGACAGCACTGGGTGCGCGGCGACGGCTGTGTGAGAGCTGCGAC 3498
DB 3282 AGCTGCTGGCTGTGACAGCACTGGGTGCGCGGCGACGGCTGTGTGAGAGCTGCGAC 3223
QY 3499 ACGAGCTGTGTGTGCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCT 3558
DB 3222 ACGAGCTGTGTGTGCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCT 3163
QY 3559 TCCGAGAGCTAGACACGGCGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3618
DB 3162 TCCGAGAGCTAGACACGGCGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACA 3103
QY 3619 CCAGCTTCGAGAGATCTTCTGAAAGTGGTGGAGAGTGTCTGCGGACACAGATATGG 3678
DB 3102 CCAGCTTCGAGAGATCTTCTGAAAGTGGTGGAGAGTGTCTGCGGACACAGATATGG 3043
QY 3679 AGGATGGCAGCTGGCGGACGACCTATGACAGGCAATTCCTGGCTAGACGTAACCTGCG 3738
DB 3042 AGGATGGCAGCTGGCGGACGACCTATGACAGGCAATTCCTGGCTAGACGTAACCTGCG 2983
QY 3739 GGCTCAAGATGCCGACACAGGAGACAGCGCTGGAGAAACGGGAAACGAGCTGGGTCAAGCCC 3798
DB 2982 GGCTCAAGATGCCGACACAGGAGACAGCGCTGGAGAAACGGGAAACGAGCTGGGTCAAGCCC 2923
QY 3799 CAGAGCTGACACAGGCTGTGGGCGACAGCGCTGGGCGGGTACAGGGCTGGGCACTGA 3858
DB 2922 CAGAGCTGACACAGGCTGTGGGCGACAGCGCTGGGCGGGTACAGGGCTGGGCACTGA 2863
QY 3859 CCGCGCAGCAGCTCCAGGCGCTGCTTCTCAGCGCTTCTGCTGGCGCGCGCAGCGCGC 3918
DB 2862 CCGCGCAGCAGCTCCAGGCGCTGCTTCTCAGCGGNGGGGGGGGGGGGGGGGGGG 2803
QY 3919 GCGGCTGTTGCGCGCAGATCGTGTGCTGCGCTCTTGTGGGCGCTGCGCTGTTCA 3978
DB 2802 GCGGCTGCT----- 2793
QY 3979 GCCTACGTGCTCTCTTTTCGGGCACTACCCGGCTCTGCGGCTCAGTCCCACTATGTACG 4038
DB 2792 ----- 2793
QY 4039 GTGCTCAGGTGCTCTTCTTCACTGAGAGCGCCCGACGGGACCTTGAGCGTCCCGGCTGC 4098
DB 2792 -----GTTTCGCGCCAGTGAGGACGCCACGAGGGACCTTGAGCGTCCCGGCTGC 2745
QY 4099 TCGAGGCGCTGCTGCAGGAGCAGGACTGGAGAGCGCCCGCAGTGCGAGCATAGCTCCACAC 4158
DB 2744 TCGAGGCGCTGCTGCAGGAGCAGGACTGGAGAGCGCCCGCAGTGCGAGCATAGCTCCACAC 2685
QY 4159 GGTTCGCGCACCAAGAGTTCTGCTGAAAGTGCGCAAGGTCTTGGCCAGTGGCAACTGGA 4218
DB 2684 GGTTCGCGCACCAAGAGTTCTGCTGAAAGTGCGCAAGGTCTTGGCCAGTGGCAACTGGA 2625
QY 4219 CCCAGAGTTCATTCGCCAGCTGCGAGTGTAGCGAGCGCGGTGCGCGGCGCTGTGTC 4278
DB 2624 CCCAGAGTTCATTCGCCAGCTGCGAGTGTAGCGAGCGCGGTGCGCGGCGCTGTGTC 2565
QY 4279 CCGAGTCCCGGCTGCGAGCTGGTGTCTCCCTCGCGCCAGGAGTGTAGCGGCTCTGGGG 4338
DB 2564 CCGAGTCCCGGCTGCGAGCTGGTGTCTCCCTCGCGCCAGGAGTGTAGCGGCTCTGGGG 2505
QY 4339 AAGTGGTTCAGAACCTGACAGCGCGGAACTGTCTGACTTCTCTGCTCAAGACCTACCGCG 4398

DB 2504 AAGTGGTTCAGAACCTGACAGCGCGGAACTGTCTGACTTCTGTCTAAGACTACCCGCG 2445
QY 4399 GCCTGTGCGCGCAGGCGCTGAAGACTAAGAGTGGGTGAATGAGGTCAAGGTACGAGGCT 4458
DB 2444 GCCTGTGCGCGCAGGCGCTGAAGACTAAGAGTGGGTGAATGAGGTCAAGGTACGAGGCT 2385
QY 4459 TCTGCTGGGGGGCGGAGACCCAGGCTTGCCTCGGGCAAGAGTGGGCGCTCAGTGG 4518
DB 2384 TCTGCTGGGGGGCGGAGACCCAGGCTTGCCTCGGGCAAGAGTGGGCGCTCAGTGG 2325
QY 4519 AGGAGTTGTGGCGCTGTGAGTCCCTGCTGCGCGGGGCGCTCGACCGTGTCTGAAAA 4578
DB 2324 AGGAGTTGTGGCGCTGTGAGTCCCTGCTGCGCGGGGCGCTCGACCGTGTCTGAAAA 2265
QY 4579 ACCTCACAGCTGGGCTCACAGCTTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACA 4638
DB 2264 ACCTCACAGCTGGGCTCACAGCTTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACA 2205
QY 4639 ACAAGGCTGGCACTTCCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCTCGTG 4698
DB 2204 ACAAGGCTGGCACTTCCATGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCTCGTG 2145
QY 4699 CTCACTGCGCCCGAGGCGCGCGCCACGCGCCACAGCATCACACACTCAACACCCCT 4758
DB 2144 CTCACTGCGCCCGAGGCGCGCGCCACGCGCCACAGCATCACACACTCAACACCCCT 2085
QY 4759 TGAACCTCACCAAGGAGCGCTGTGAGGCTGCACTGATGGCTCTCTCGGTGGACGCTCC 4818
DB 2084 TGAACCTCACCAAGGAGCGCTGTGAGGCTGCACTGATGGCTCTCTCGGTGGACGCTCC 2025
QY 4819 TCGTCTCCATCTGTGTGTCTTTGCCATGTCTTTGTCGGCGCAGCTTCACTCTTGTTC 4878
DB 2024 TCGTCTCCATCTGTGTGTCTTTGCCATGTCTTTGTCGGCGCAGCTTCACTCTTGTTC 1965
QY 4879 TCATTGAGGAGCGAGTCAACCGAGCCAAAGCACTGCACTCATGGGGGCGCTGTCCCCCA 4938
DB 1964 TCATTGAGGAGCGAGTCAACCGAGCCAAAGCACTGCACTCATGGGGGCGCTGTCCCCCA 1905
QY 4939 CCCTCTACTGGCTTGGCAACTTTCTCTGGAGATGTGTAACTACTTGGTGCAGCATGCA 4998
DB 1904 CCCTCTACTGGCTTGGCAACTTTCTCTGGAGATGTGTAACTACTTGGTGCAGCATGCA 1845
QY 4999 TCGTGTGCTCATCTTTCTGGCTTCCAGCAGAGGCGATATGTGCGCCCTGCAACCTGCG 5058
DB 1844 TCGTGTGCTCATCTTTCTGGCTTCCAGCAGAGGCGATATGTGCGCCCTGCAACCTGCG 1785
QY 5059 CTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATCACACCGCTCATGTACCCAG 5118
DB 1784 CTGCTCTCTGCTGTGCTACTACTGTATGGCTGGTGCATCACACCGCTCATGTACCCAG 1725
QY 5119 CCTCTCTCTTCTTCTCGTGGCGAGCAGCGCTATGTGGTGCCTCACTGCAATAAACCTCT 5178
DB 1724 CCTCTCTCTTCTTCTCGTGGCGAGCAGCGCTATGTGGTGCCTCACTGCAATAAACCTCT 1665
QY 5179 TTATTGGCATCAATGGAAGATGGCCACCTTTGTGCTTGTAGCTCTTCTCATCAGAGC 5238
DB 1664 TTATTGGCATCAATGGAAGATGGCCACCTTTGTGCTTGTAGCTCTTCTCATCAGAGC 1605
QY 5239 TGCAGGAGGTGAGCCGATCTTGAAACAGGCTTTCTTATCTTCCCCCACTTCTGCTGG 5298
DB 1604 TGCAGGAGGTGAGCCGATCTTGAAACAGGCTTTCTTATCTTCCCCCACTTCTGCTGG 1545
QY 5299 GCGGGGGCTCATTTGACATGCTGGGAAACGAGGCTATGTGGTGCCTTTGAGGCGCTGG 5358
DB 1544 GCGGGGGCTCATTTGACATGCTGGGAAACGAGGCTATGTGGTGCCTTTGAGGCGCTGG 1485
QY 5359 GAGACGCGATTCAGTCACTCCCTGCTGGAGGTGCTGGCAAGAACCTCTTTGGCCA 5418
DB 1484 GAGACGCGATTCAGTCACTCCCTGCTGGAGGTGCTGGCAAGAACCTCTTTGGCCA 1425
QY 5419 TGGTGTATACAGGGGCGCTCTTCTTCTTCACTACTGCTGACGACCGAAGCCAAAC 5478

CC disorders, dyslipidaemia, epilepsy, diseases related to abnormal calcium
CC flux, coronary artery disease, Tangier's disease, familial high-density
CC lipoprotein deficiency, atherosclerosis, diabetes, fatty liver disease,
CC insulin resistance, obesity, alcoholism, retinal degeneration,
CC hypertension and vascular disease. The sequences are also used in drug
CC screening assays
XX

SQ Sequence 5669 BP; 965 A; 1843 C; 1745 G; 1116 T; 0 U; 0 Other;

Query Match 80.2%; Score 5458.2; DB 4; Length 5669;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5460; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	210	ATGGCCTTCTGGACACAGCTGATGCTGCTCTGGAAGATTTTCATGTATCGCGGAGA	269
Db	1	ATGGCCTTCTGGACACAGCTGATGCTGCTCTGGAAGATTTTCATGTATCGCGGAGA	60
Qy	270	CAGCGGTCCAGCTCCTGGTCGAATTTGCTGGCCCTCTTCCTCTTCTTATCTGTGTG	329
Db	61	CAGCGGTCCAGCTCCTGGTCGAATTTGCTGGCCCTCTTCCTCTTCTTATCTGTGTG	120
Qy	330	GCTGTTCCGCACTCCACCCGCGCTGGAGCACCATGAATGCCACTTCCCAAAACAAGCA	389
Db	121	GCTGTTCCGCACTCCACCCGCGCTGGAGCACCATGAATGCCACTTCCCAAAACAAGCA	180
Qy	390	CTGCCATCGCGGGGACCGTGCCTCGCTCCAGGGTCTCATCTGTAATGTGAACAACACC	449
Db	181	CTGCCATCGCGGGGACCGTGCCTCGCTCCAGGGTCTCATCTGTAATGTGAACAACACC	240
Qy	450	TGCTTTCCGCACTGACACCGGGGAGAGGCCCGGGCGCTGAGCAATTCACAGACTCC	509
Db	241	TGCTTTCCGCACTGACACCGGGGAGAGGCCCGGGCGCTGAGCAATTCACAGACTCC	300
Qy	510	CTGGTCTCCCGCTGTAGCGATGCCCGCACTGTGCTGGAGGGCCAGTGCCCAACAGG	569
Db	301	CTGGTCTCCCGCTGTAGCGATGCCCGCACTGTGCTGGAGGGCCAGTGCCCAACAGG	360
Qy	570	ACGTGGCTGGCTAGGGAAGCTGATCGCCAGCTGAGGGTGCACGAGCAACGGCCACG	629
Db	361	ACGTGGCTGGCTAGGGAAGCTGATCGCCAGCTGAGGGTGCACGAGCAACGGCCACG	420
Qy	630	CCTCAACCAACAAAGCACTCTCACTGAAACACCCATGTGGATGTCCGCGAGCTGTG	689
Db	421	CCTCAACCAACAAAGCACTCTCACTGAAACACCCATGTGGATGTCCGCGAGCTGTG	480
Qy	690	ACGTCACTGTCCGACAGGAATCCCTGGGTTGGCACTGGGCCAAGCCAGAGGCCCTTG	749
Db	481	ACGTCACTGTCCGACAGGAATCCCTGGGTTGGCACTGGGCCAAGCCAGAGGCCCTTG	540
Qy	750	CACAGCTTGTGGAGGCGCTGAGGACCTGGCCACAGGAGCTCCTGGCGCTGCGAGCCTG	809
Db	541	CACAGCTTGTGGAGGCGCTGAGGACCTGGCCACAGGAGCTCCTGGCGCTGCGAGCCTG	600
Qy	810	GTGAGCTTCGGGCACTGCTCAGAGACCCCGAGGGACAGCGGCCCTCGAGTTGCTG	869
Db	601	GTGAGCTTCGGGCACTGCTCAGAGACCCCGAGGGACAGCGGCCCTCGAGTTGCTG	660
Qy	870	TCAGAGGCCCTTGCAAGTGTACGGGACCTAGCAGCACAGTGGGCCCTCCTCAACTGG	929
Db	661	TCAGAGGCCCTTGCAAGTGTACGGGACCTAGCAGCACAGTGGGCCCTCCTCAACTGG	720
Qy	930	TACGAGGCTAGTGACCTGATGGAGCTGTGGGGACGAGATCCGCTGCCAGAC	989
Db	721	TACGAGGCTAGTGACCTGATGGAGCTGTGGGGACGAGATCCGCTGCCAGAC	780
Qy	990	AGCAGCTGAGCCCGCTGCTCGAGAGCTGATTTGAGCCCTGGACAGCAACCGCTGTCC	1049
Db	781	AGCAGCTGAGCCCGCTGCTCGAGAGCTGATTTGAGCCCTGGACAGCAACCGCTGTCC	840
Qy	1050	CGCTGTCTCGAGACGCTGAAGCCTCTGATCTCGGGAAGCTACTCTTTGCAACAGAT	1109
Db	841	CGCTGTCTCGAGACGCTGAAGCCTCTGATCTCGGGAAGCTACTCTTTGCAACAGAT	900

Qy	1110	ACACCTTTTACC CGAAGCTCATGSCCAGAGTGAAACCGAGACCTTCGAGAGCTCACCCCTG	1169
Db	901	ACACCTTTTACC CGAAGCTCATGSCCAGAGTGAAACCGAGACCTTCGAGAGCTCACCCCTG	960
Qy	1170	CTGAGGGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTACCTTTCATGAAC	1229
Db	961	CTGAGGGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTACCTTTCATGAAC	1020
Qy	1230	GACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGACAGATGCAGATGAAGAAGAGG	1289
Db	1021	GACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGACAGATGCAGATGAAGAAGAGG	1080
Qy	1290	CAGCCAGACCTGAGGCGGGACACATGAGAGCCCTGCGATCCTTTCTGGAACCTCGG	1349
Db	1081	CAGCCAGACCTGAGGCGGGACACATGAGAGCCCTGCGATCCTTTCTGGAACCTCGG	1140
Qy	1350	AGCGTGGCTACAGCTGSCAGAGCGCACACCTGATGTGGGGCACCTGGTGGGACGCTG	1409
Db	1141	AGCGTGGCTACAGCTGSCAGAGCGCACACCTGATGTGGGGCACCTGGTGGGACGCTG	1200
Qy	1410	GGCCAGTGCAGAGTGCCTGCTTGGACAAGCTGGAGGCGGACACCTCAGAGGACGAC	1469
Db	1201	GGCCAGTGCAGAGTGCCTGCTTGGACAAGCTGGAGGCGGACACCTCAGAGGACGAC	1260
Qy	1470	CTGGTGTCCGCGGCCCTGCAACTGCTCCGCGAAACATCGATTTCTGGGCGGGCTCTTTC	1529
Db	1261	CTGGTGTCCGCGGCCCTGCAACTGCTCCGCGAAACATCGATTTCTGGGCGGGCTCTTTC	1320
Qy	1530	TTGGGACCTGAGGACTCTTTCAGACCCACAGAGCAACCAACCCAGACCTTGGGCCCGGC	1589
Db	1321	TTGGGACCTGAGGACTCTTTCAGACCCACAGAGCAACCAACCCAGACCTTGGGCCCGGC	1380
Qy	1590	CACGTGGCATCAAAATCCGCATGGACATTCACGTTGTCAGAGGACCAATAAGATCAGG	1649
Db	1381	CACGTGGCATCAAAATCCGCATGGACATTCACGTTGTCAGAGGACCAATAAGATCAGG	1440
Qy	1650	GACAGTTTTGGGACCCCTGGCCAGCGCGACCCCTCGACCGACCTCGCTACGCTGTGG	1709
Db	1441	GACAGTTTTGGGACCCCTGGCCAGCGCGACCCCTCGACCGACCTCGCTACGCTGTGG	1500
Qy	1710	GGCGGCTTCTGTTACTGTCAAAGACTGTGTGAGCGGTGACCGCTCCGCTGTCTAGCGGC	1769
Db	1501	GGCGGCTTCTGTTACTGTCAAAGACTGTGTGAGCGGTGACCGCTCCGCTGTCTAGCGGC	1560
Qy	1770	GCCAAACCCCGGGCGGCTCTACCTGCAGCAGATGCCCTATCGTGTCTATGTGACCGAC	1829
Db	1561	GCCAAACCCCGGGCGGCTCTACCTGCAGCAGATGCCCTATCGTGTCTATGTGACCGAC	1620
Qy	1830	GTGTTCTCTGCTGTGAGCGGCTCGCTGCCGCTCTTCTGACGCTCGGCTGATCTAC	1889
Db	1621	GTGTTCTCTGCTGTGAGCGGCTCGCTGCCGCTCTTCTGACGCTCGGCTGATCTAC	1680
Qy	1890	TCCGTGACACTGA CAGTGAAGCGCGTGTGCGGAGAAAGAGACGCGCTCGCGGACACC	1949
Db	1681	TCCGTGACACTGA CAGTGAAGCGCGTGTGCGGAGAAAGAGACGCGCTCGCGGACACC	1740
Qy	1950	ATGCGGCCATGSGGCTCAGCGCGCGTGTCTGCGCTAGGCTGGTCTCCTCAGCTGCTC	2009
Db	1741	ATGCGGCCATGSGGCTCAGCGCGCGTGTCTGCGCTAGGCTGGTCTCCTCAGCTGCTC	1800
Qy	2010	GGGCGCTTCTGCTCAGCGCGCGCTGCTGCTTCTGCTGCTCAAGCTGGGGACATCTCTC	2069
Db	1801	GGGCGCTTCTGCTCAGCGCGCGCTGCTGCTTCTGCTGCTCAAGCTGGGGACATCTCTC	1860
Qy	2070	CCCTACAGCAACCCGGGCGTGTCTTCTGTTTGTGGCAGCTTTCGCGGTGGCCAAGGTG	2129
Db	1861	CCCTACAGCAACCCGGGCGTGTCTTCTGTTTGTGGCAGCTTTCGCGGTGGCCAAGGTG	1920
Qy	2130	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGCGGCTGCGGC	2189
Db	1921	ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGCGGCTGCGGC	1980
Qy	2190	GGCCTGGCCTACTTCTCCTCTACCTGCCCTACGTCGCTGTGTGTGGCTTGGCGGACCGG	2249

1981 DB |||||GGCCTG|GCTTCT|CCCTCT|ATCTG|CCCTCT|ATCGT|GCTGTGTGTG|GCTTGG|CGGGAC|CGG 2040
2250 QY CTGCCG|CGGGTGG|CGCGTGG|CGGAGC|CTGCTGTG|CGCCGCTGG|CGCTTTCGG|CTTCGGC 2309
2041 DB CTGCCG|CGGGTGG|CGCGTGG|CGGAGC|CTGCTGTG|CGCCGCTGG|CGCTTTCGG|CTTCGGC 2100
2310 QY TGGAGAG|CCTGGCTCT|GCTGG|AGGAG|CAGGGG|CGGCGAG|TGGGCA|CAACGTGGGC 2369
2101 DB TGGAGAG|CCTGGCTCT|GCTGG|AGGAG|CAGGGG|CGGCGAG|TGGGCA|CAACGTGGGC 2160
2370 QY ACCGGC|CTACGG|CCTG|GAGC|GCTTT|CAGC|CTGG|GCCCA|GGTCT|CTGG|CCTT|CTGCT|GCTGGAC 2429
2161 DB ACCGGC|CTACGG|CCTG|GAGC|GCTTT|CAGC|CTGG|GCCCA|GGTCT|CTGG|CCTT|CTGCT|GCTGGAC 2220
2430 QY GCGGCG|CTACGG|CCTG|GAGC|GCTTT|CAGC|CTGG|TACT|CTGGA|AGCTGT|GCTCC|AGGCGAGT|ACGGG 2489
2221 DB GCGGCG|CTACGG|CCTG|GAGC|GCTTT|CAGC|CTGG|TACT|CTGGA|AGCTGT|GCTCC|AGGCGAGT|ACGGG 2280
2490 QY ATCCCTG|AACCAT|GGAATTT|CTTTT|CGGAGG|AGCTACT|GCTGG|AGACCT|CGGCC|CCCC 2549
2281 DB ATCCCTG|AACCAT|GGAATTT|CTTTT|CGGAGG|AGCTACT|GCTGG|AGACCT|CGGCC|CCCC 2340
2550 QY AAGAGTC|CAGCC|CTT|GCCCA|CCCCG|CTGGA|CCCCA|AAGGTG|CTGTG|TAGA|AGGCA|CCG 2609
2341 DB AAGAGTC|CAGCC|CTT|GCCCA|CCCCG|CTGGA|CCCCA|AAGGTG|CTGTG|TAGA|AGGCA|CCG 2400
2610 QY CCGGCG|CTGAGT|CTGCG|GCTA|TCG|TTG|CGAGC|CTGGA|AAGCG|CTTT|CTGGA|AGCGCG 2669
2401 DB CCGGCG|CTGAGT|CTGCG|GCTA|TCG|TTG|CGAGC|CTGGA|AAGCG|CTTT|CTGGA|AGCGCG 2460
2670 QY CAGCAG|CCTG|CGGGG|CTCAG|CTGGA|CTT|CTA|CCAGG|GCCAC|ATC|ACG|CGCTT|CCTG 2729
2461 DB CAGCAG|CCTG|CGGGG|CTCAG|CTGGA|CTT|CTA|CCAGG|GCCAC|ATC|ACG|CGCTT|CCTG 2520
2730 QY GGCACAA|CGGGG|CGGCA|AGAC|CAC|CAC|CTG|TCCAT|TTGAGT|GCGCT|CTT|CCAC|CC 2789
2521 DB GGCACAA|CGGGG|CGGCA|AGAC|CCA|CCA|CTG|TCCAT|TTGAGT|GCGCT|CTT|CCAC|CC 2580
2790 QY AGTGGT|GCTG|CTG|CTT|CAT|CTCG|GGCC|ACG|ATG|CTCG|CTC|ACG|ATG|GCG|CGCAT|CCGG 2849
2581 DB AGTGGT|GCTG|CTG|CTT|CAT|CTCG|GGCC|ACG|ATG|CTCG|CTC|ACG|ATG|GCG|CGCAT|CCGG 2640
2850 QY CCCACCT|GGGCTG|CTGCT|CAGT|ACAA|CGT|GCTG|TTT|GACAT|GCTG|ACCG|TGGAG|CGAG 2909
2641 DB CCCACCT|GGGCTG|CTGCT|CAGT|ACAA|CGT|GCTG|TTT|GACAT|GCTG|ACCG|TGGAG|CGAG 2700
2910 QY CAGCTCT|GGTTC|ATGGG|CGGCTG|AAGG|GCTG|AGT|GCG|CTGT|AGT|GGG|CGCC|CGAG|CGAG 2969
2701 DB CAGCTCT|GGTTC|ATGGG|CGGCTG|AAGG|GCTG|AGT|GCG|CTGT|AGT|GGG|CGCC|CGAG|CGAG 2760
2970 QY GACCGT|CTGTC|GAGG|ATG|GGG|GCTG|GCT|C|CAAG|CAGAGT|GTG|CAGAT|CTG|CGCA|CCTC 3029
2761 DB GACCGT|CTGTC|GAGG|ATG|GGG|GCTG|GCT|C|CAAG|CAGAGT|GTG|CAGAT|CTG|CGCA|CCTC 2820
3030 QY TCTGTGG|GATG|CAAC|CGGA|AGCTG|TCCG|TGGG|CAAT|TGC|CTTT|GTG|GGG|CGGCT|CCCA|GTT 3089
2821 DB TCTGTGG|GATG|CAAC|CGGA|AGCTG|TCCG|TGGG|CAAT|TGC|CTTT|GTG|GGG|CGGCT|CCCA|GTT 2880
3090 QY GTTATCT|GTGAC|GAGC|CTA|CGG|CTGGG|TGGA|TCTG|CTT|CCCG|CGCGG|TATTT|GGGAG 3149
2881 DB GTTATCT|GTGAC|GAGC|CTA|CGG|CTGGG|TGGA|TCTG|CTT|CCCG|CGCGG|TATTT|GGGAG 2940
3150 QY CTGCTGT|CTAAAT|CCGA|AGG|TGGCA|CGCTG|ATCT|CTT|CCA|CCCA|CCCTG|GATGAG 3209
2941 DB CTGCTGT|CTAAAT|CCGA|AGG|TGGCA|CGCTG|ATCT|CTT|CCA|CCCA|CCCTG|GATGAG 3000
3210 QY GCAGAGT|GCTGG|GAGAC|CGTGT|GGCTGT|GGTGG|CAGG|TGG|CGCTT|GCTGCTG|CTGGC 3269
3001 DB GCAGAGT|GCTGG|GAGAC|CGTGT|GGCTGT|GGTGG|CAGG|TGG|CGCTT|GCTGCTG|CTGGC 3060
3270 QY TCCCACT|CTTCT|CTG|CGCG|CTCA|CTGGG|CTCGG|GCTACT|GAC|CGCTG|GTTG|GGA|GGCC 3329

3061 DB TCCCACT|CTTCT|CTG|CGCG|CTCA|CTGGG|CTCCG|CTACTA|CTGAGC|GTGGTGA|AGGCC 3120
3330 QY CGCCTC|CCCCTG|ACAC|CAATG|AAGG|CTGAC|ACTG|ATGATG|AGGCG|CAGTGTGGA|CACC 3389
3121 DB CGCCTC|CCCCTG|ACAC|CAATG|AAGG|CTGAC|ACTG|ATGATG|AGGCG|CAGTGTGGA|CACC 3180
3390 QY AGCAGG|AAAAG|AAGTGC|AGCC|AGGG|CAGCAG|AGTGTGG|CACTCT|CTCAG|CTGTG|GGCC 3449
3181 DB AGCAGG|AAAAG|AAGTGC|AGCC|AGGG|CAGCAG|AGTGTGG|CACTCT|CTCAG|CTGTG|GGCC 3240
3450 QY CTGGTA|CAGCA|CTGGGTG|CCCG|GGGCA|CGGCTGGTGG|AGAGCTGCCA|CAGAGCTGGTG 3509
3241 DB CTGGTA|CAGCA|CTGGGTG|CCCG|GGGCA|CGGCTGGTGG|AGAGCTGCCA|CAGAGCTGGTG 3300
3510 QY CTGGTCT|GCCCTT|ACAGG|GTGCC|ATGACGGG|CAGCTT|CGCCAC|ACTT|TCCGAGAGCTA 3569
3301 DB CTGGTCT|GCCCTT|ACAGG|GTGCC|ATGACGGG|CAGCTT|CGCCAC|ACTT|TCCGAGAGCTA 3360
3570 QY GACACG|CGGCTGG|CGGAGCTG|AGGCTCA|CTGGCTAG|GGGATCTT|CCGACAC|CCAGCCTCGAG 3629
3361 DB GACACG|CGGCTGG|CGGAGCTG|AGGCTCA|CTGGCTAG|GGGATCTT|CCGACAC|CCAGCCTCGAG 3420
3630 QY GAGATCTT|CTGAA|GGTGTGG|AGGAGTGTG|TCGGA|CA|CAGATATG|AGGATGGG|CAGC 3689
3421 DB GAGATCTT|CTGAA|GGTGTGG|AGGAGTGTG|TCGGA|CA|CAGATATG|AGGATGGG|CAGC 3480
3690 QY TGGGGC|CAGCA|CCTATG|CA|CAGGCA|TTGCTGGG|CTTAGA|CGTAA|CCCTCGG|GCTCA|GATG 3749
3481 DB TGGGGC|CAGCA|CCTATG|CA|CAGGCA|TTGCTGGG|CTTAGA|CGTAA|CCCTCGG|GCTCA|GATG 3540
3750 QY CCGCCA|CAGGAG|CAGCGCTG|GAGAA|CGGGG|AAC|CAGCTTGG|GGTCAG|GCC|CCAGAGA|CTGAC 3809
3541 DB CCGCCA|CAGGAG|CAGCGCTG|GAGAA|CGGGG|AAC|CAGCTTGG|GGTCAG|GCC|CCAGAGA|CTGAC 3600
3810 QY CAGGGT|CTTGGG|CAGAC|CGCGTGG|CGCCGGG|TACAGGG|CTGGGCA|CTGA|CCG|CCAGCAG 3869
3601 DB CAGGGT|CTTGGG|CAGAC|CGCGTGG|CGCCGGG|TACAGGG|CTGGGCA|CTGA|CCG|CCAGCAG 3660
3870 QY CTCAGG|CCCTGCTT|CTCAAG|CGCTTTCG|TTGCTG|TCCG|CGCGAG|CGCGCGG|CGCTG|TTC 3929
3661 DB CTCAGG|CCCTGCTT|CTCAAG|CGCTTTCG|TTGCTG|TCCG|CGCGAG|CGCGCGG|CGCTG|TTC 3720
3930 QY GCCAGAT|CGTGTG|CTGCTG|CCCTCT|TTTGGG|CGTGG|CCCTCG|TGTTCAG|CCCTCAT|CGTG 3989
3721 DB GCCAGAT|CGTGTG|CTGCTG|CCCTCT|TTTGGG|CGTGG|CCCTCG|TGTTCAG|CCCTCAT|CGTG 3780
3990 QY CTTCTTT|CGGGCA|CTA|CCCGG|CTGCGG|CTCAGT|CCCA|CATATG|A|CGTGTCT|CAGGTG 4049
3781 DB CTTCTTT|CGGGCA|CTA|CCCGG|CTGCGG|CTCAGT|CCCA|CATATG|A|CGTGTCT|CAGGTG 3840
4050 QY TCCCTT|CTTCA|GTAGAG|ACG|CCCC|CAGGGG|ACCTTGG|ACG|TCCCGG|CGCTG|CTCGAGG|CGCTG 4109
3841 DB TCCCTT|CTTCA|GTAGAG|ACG|CCCC|CAGGGG|ACCTTGG|ACG|TCCCGG|CGCTG|CTCGAGG|CGCTG 3900
4110 QY CTSACG|AGGSCA|GACTG|GAGGAG|CGCC|CAGTGCAG|CATAGCTCC|CACAGG|TTCTCGGCA 4169
3901 DB CTSACG|AGGSCA|GACTG|GAGGAG|CGCC|CAGTGCAG|CATAGCTCC|CACAGG|TTCTCGGCA 3960
4170 QY CCAGAA|GTTCCT|CTGAA|GTGG|CCAA|AGGCTTGG|CCAGTGG|GCAACTGGG|AATCTGGA|CCCCAGAGTCT 4229
3961 DB CCAGAA|GTTCCT|CTGAA|GTGG|CCAA|AGGCTTGG|CCAGTGG|GCAACTGGG|AATCTGGA|CCCCAGAGTCT 4020
4230 QY CAGTCC|CGAGC|CTG|CCAGTGTAG|CCAG|CGG|GTGCC|GGG|CGCTG|TCCCGG|ACG|TCCCG 4289
4021 DB CAGTCC|CGAGC|CTG|CCAGTGTAG|CCAG|CGG|GTGCC|GGG|CGCTG|TCCCGG|ACG|TCCCG 4080
4290 QY GCTGCA|GCTGTG|GCTCC|CTCG|CGCC|CAGG|CAGTGCAG|CGGCTCTGGG|AAGTGGTTCAG 4349
4081 DB GCTGCA|GCTGTG|GCTCC|CTCG|CGCC|CAGG|CAGTGCAG|CGGCTCTGGG|AAGTGGTTCAG 4140
4350 QY AACCTG|CAGG|CGG|CAAC|CTGCTG|ACTTCTGGTCA|AGAC|CTTAC|CGCG|CGCTGGTGGC 4409
4141 DB AACCTG|CAGG|CGG|CAAC|CTGCTG|ACTTCTGGTCA|AGAC|CTTAC|CGCG|CGCTGGTGGC 4200

Db	751	CCCCGGAATCCCTGGGGTTGGCACTGGGCCAAGCCAGAGCCCTTGGCAACGTTGTTG	810
Qy	762	GAGGCGCTGAGACCTGGGCCCAAGAGACTCCTGGCGTGGCGAGCCTGGTGGAGCTTCGG	821
Db	811	GAGGCGCTGAGACCTGGGCCCAAGAGACTCCTGGCGTGGCGAGCCTGGTGGAGCTTCGG	870
Qy	822	GCACTGCTGAGAGACCCCGAGGGACAGCGGCCCTTGGAGTTGCTGTACAGAGGCCCTC	881
Db	871	GCACTGCTGAGAGACCCCGAGGGACAGCGGCCCTTGGAGTTGCTGTACAGAGGCCCTC	930
Qy	882	TGCAGTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTTCCCTCAACTGGTACAGAGGCTAGT	941
Db	931	TGCAGTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTTCCCTCAACTGGTACAGAGGCTAGT	990
Qy	942	GACCTGATGAGCTGTGTGGGCGAGGACCAAGATCCGCGCTGCCAGACAGACGCTGAGC	1001
Db	991	GACCTGATGAGCTGTGTGGGCGAGGACCAAGATCCGCGCTGCCAGACAGACGCTGAGC	1050
Qy	1002	CCGCGCTGCTCGAGCTGATTGGAGCCCTGGACAGCCACCGCTGTCCCGCCTGCTTGG	1061
Db	1051	CCGCGCTGCTCGAGCTGATTGGAGCCCTGGACAGCCACCGCTGTCCCGCCTGCTTGG	1110
Qy	1062	AGACGCTGAAGCTCTGTGATCCTCGGGAAGCTACTCTTTGCACAGATACACCTTTTACC	1121
Db	1111	AGACGCTGAAGCTCTGTGATCCTCGGGAAGCTACTCTTTGCACAGATACACCTTTTACC	1170
Qy	1122	CGAAGCTCATGSCCCAGGTGAACCGGACCTTCGAGAGCTCACCTTCATGAACGACGTTCCAT	1181
Db	1171	CGAAGCTCATGSCCCAGGTGAACCGGACCTTCGAGAGCTCACCTTCATGAACGACGTTCCAT	1230
Qy	1182	CGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAACGACGTTCCAT	1241
Db	1231	CGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTCATGAACGACGTTCCAT	1290
Qy	1242	GTGSCCATGCTGACGGCTCCTGCAGATGCAGATGAAGGAAGAGGACGCCAGACCT	1301
Db	1291	GTGSCCATGCTGACGGCTCCTGCAGATGCAGATGAAGGAAGAGGACGCCAGACCT	1350
Qy	1302	GGAGGCGGGACACATGAGGCGCTCGATCCTTTCTGGACCTCGGAGCGGTGCTAC	1361
Db	1351	GGAGGCGGGACACATGAGGCGCTCGATCCTTTCTGGACCTCGGAGCGGTGCTAC	1410
Qy	1362	AGTGGCAGACGCAACGCTGATGTGGGGCACTGGTGGGACGCTGGGCCAGGTGACG	1421
Db	1411	AGTGGCAGACGCAACGCTGATGTGGGGCACTGGTGGGACGCTGGGCCAGGTGACG	1470
Qy	1422	GAGTGTCTGCTTGGACAAGCTGGAGCGGCCACCTTCAGAGCAGCCCTGGTGTGCGGG	1481
Db	1471	GAGTGTCTGCTTGGACAAGCTGGAGCGGCCACCTTCAGAGCAGCCCTGGTGTGCGGG	1530
Qy	1482	GCCCTGCACTGTGCGGGAACATCGATTCTGGGCGGCGCTCGTCTTTTGGGACCTGAG	1541
Db	1531	GCCCTGCACTGTGCGGGAACATCGATTCTGGGCGGCGCTCGTCTTTTGGGACCTGAG	1590
Qy	1542	GACTCTTCAGACCCGACAGACGCCAACCCGACGCTGGGCCCGCGGCACTGGCGCATC	1601
Db	1591	GACTCTTCAGACCCGACAGACGCCAACCCGACGCTGGGCCCGCGGCACTGGCGCATC	1650
Qy	1602	AAATCCGATGACATTTGACGTGTTCAGGAGACCAATGAAGATCAGGGACAGGTTTGG	1661
Db	1651	AAATCCGATGACATTTGACGTGTTCAGGAGACCAATGAAGATCAGGGACAGGTTTGG	1710
Qy	1662	GACCTGGCCGACGCGGACCCCTGAACGACCTGGCTACGCTGGGCGGCGCTTCGTTG	1721
Db	1711	GACCTGGCCGACGCGGACCCCTGAACGACCTGGCTACGCTGGGCGGCGCTTCGTTG	1770
Qy	1722	TACCTGCAACGCTGTGGAGCGTGGAGCCGCTCCGCGTGTCTACGCGCGCCCAACCCCGG	1781
Db	1771	TACCTGCAACGCTGTGGAGCGTGGAGCCGCTCCGCGTGTCTACGCGCGCCCAACCCCGG	1830
Qy	1782	GCCGGCTCTACCTGACGAGATGCCCTATCCGCTGTATGTGGAACGAGCTTCCTGCGT	1841

Db	1831	GCCGGCTCTACCTGACGAGATGCCCTATCCGTGCTATGTGGAACGAGCTTCCTGCGT	1890
Qy	1842	GTGCTGAGCCGGTGGCTGCGCTCTTCTGAGACGCTGGCTGGATCTACTCCGTGACACTG	1901
Db	1891	GTGCTGAGCCGGTGGCTGCGCTCTTCTGAGACGCTGGCTGGATCTACTCCGTGACACTG	1950
Qy	1902	ACAGTGAAGCGCTGTGTGGGAGAGAGACGCGCTGCGGACACCATGCGCGGCATG	1961
Db	1951	ACAGTGAAGCGCTGTGTGGGAGAGAGACGCGCTGCGGACACCATGCGCGGCATG	2010
Qy	1962	GGGCTCAGCGCGCGCTGTCTGGCTAGGCTGGTCTTCTCAGCTGCTCGGCGCTTCCTG	2021
Db	2011	GGGCTCAGCGCGCGCTGTCTGGCTAGGCTGGTCTTCTCAGCTGCTCGGCGCTTCCTG	2070
Qy	2022	CTCAGCGCGCGCTGTCTGGCTAGGCTGGTCTTCTCAGCTGCGGGGACATCTTCCCTACAGGCAC	2081
Db	2071	CTCAGCGCGCGCTGTCTGGCTAGGCTGGTCTTCTCAGCTGCGGGGACATCTTCCCTACAGGCAC	2130
Qy	2082	CGGGCGTGTCTTCTGTCTTGGCAGCTTCCGCGTGGCCACGCTGACCCAGAGCTTC	2141
Db	2131	CGGGCGTGTCTTCTGTCTTGGCAGCTTCCGCGTGGCCACGCTGACCCAGAGCTTC	2190
Qy	2142	CTGCTCAGCGCTTCTTCTCCGCGCAACCTGGCTGCGGCTGCGCGGCTGGGCTAC	2201
Db	2191	CTGCTCAGCGCTTCTTCTCCGCGCAACCTGGCTGCGGCTGCGCGGCTGGGCTAC	2250
Qy	2202	TTCTCCCTCTACCTGCGCTTACGCTGTGTGTGGCTTGGGGGACCGGCTGCCCGGGT	2261
Db	2251	TTCTCCCTCTACCTGCGCTTACGCTGTGTGTGGCTTGGGGGACCGGCTGCCCGGGT	2310
Qy	2262	GGCGCGTGGCGCGGAGCTGTCTGCGCGCTGGCTTGGCTTGGCGGAGCCGCTGCGGAGCGCTG	2321
Db	2311	GGCGCGTGGCGCGGAGCTGTCTGCGCGCTGGCTTGGCTTGGCGGAGCCGCTGCGGAGCGCTG	2370
Qy	2322	GCTCTCTGGAGGAGCAGGGCGAGGCGCGGACAAAGCTGGGCAACCGGCGCTACG	2381
Db	2371	GCTCTCTGGAGGAGCAGGGCGAGGCGCGGACAAAGCTGGGCAACCGGCGCTACG	2430
Qy	2382	GCAGAGCTCTTACGCTGCGGCCAGGCTCTGGCTTCTGGCTTCTGGCTGAGCGGGCGCTTAC	2441
Db	2431	GCAGAGCTCTTACGCTGCGGCCAGGCTCTGGCTTCTGGCTTCTGGCTGAGCGGGCGCTTAC	2490
Qy	2442	GGCTCGCCACCTGTGTACCTGGAAAGCTGTGTGCCAGGCGCAGTAACGCGATCCCTGAAACCA	2501
Db	2491	GGCTCGCCACCTGTGTACCTGGAAAGCTGTGTGCCAGGCGCAGTAACGCGATCCCTGAAACCA	2550
Qy	2502	TGGAATTTTCTTTTTCGAGGAGCTACTGCTGCGGACCTCGGCGCCCGGAGAGTCCAGCC	2561
Db	2551	TGGAATTTTCTTTTTCGAGGAGCTACTGCTGCGGACCTCGGCGCCCGGAGAGTCCAGCC	2610
Qy	2562	CCTTGGCCCAACCCCGCTGGACCCAAAGGTGCTGGTAGAGAGGCAACGCGCGCGCTGAGT	2621
Db	2611	CCTTGGCCCAACCCCGCTGGACCCAAAGGTGCTGGTAGAGAGGCAACGCGCGCGCTGAGT	2670
Qy	2622	CCTTGGGATACCTGCTGCGAGCTTGGAGAGCGCTTCTTGGAGAGCGCGCAGAGCGGCTG	2681
Db	2671	CCTTGGGATACCTGCTGCGAGCTTGGAGAGCGCTTCTTGGAGAGCGCGCAGAGCGGCTG	2730
Qy	2682	CGGGGCTCAGCTTGGACCTTCTACAGGCGCAACATCACCGCTTCTTGGGCGCAACGCGG	2741
Db	2731	CGGGGCTCAGCTTGGACCTTCTACAGGCGCGCAACATCACCGCTTCTTGGGCGCAACGCGG	2790
Qy	2742	GCGGGAAGACCAACACCTGCTTCTTGGAGTGGCTTCTTCCCAACGCTGCTGCTCT	2801
Db	2791	GCGGGAAGACCAACACCTGCTTCTTGGAGTGGCTTCTTCCCAACGCTGCTGCTCT	2850
Qy	2802	GCTTCATCTGGGCGACGAGCTCCGCTCAGCATGGCGGCGCATCGGCGGCGGCTGGG	2861
Db	2851	GCTTCATCTGGGCGCGACGAGCTCCGCTCAGCATGGCGGCGCATCGGCGGCGGCTGGG	2910
Qy	2862	GTCTGTCTCAGTAAACGCTGTGTTTGAATGCTGACCGTGGACGAGACGCTGCTGCTTC	2921
Db	2911	GTCTGTCTCAGTAAACGCTGTGTTTGAATGCTGACCGTGGACGAGACGCTGCTGCTTC	2970

QY 2922 TATGGGCGGCTGAAGGCTCTGATGCCGCTGTAGTGGGCCCCGAGACGACCGTCTGCTG 2981
DB |||||
DB 2971 TATGGGCGGCTGAAGGCTCTGATGCCGCTGTAGTGGGCCCCGAGACGACCGTCTGCTG 3030
QY 2982 CAGGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGACTCGGCCACTCTCTGTGGGATG 3041
DB |||||
DB 3031 CAGGATGTGGGCTGGTCTCCAGCAGAGTGTGCAGACTCGGCCACTCTCTGTGGGATG 3090
QY 3042 CAACGGAAGCTGTCCGCTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTTGTTATCTCGAC 3101
DB |||||
DB 3091 CAACGGAAGCTGTCCGCTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTTGTTATCTCGAC 3150
QY 3102 GAGCCTACGGCTGGGCTGATCCTGCTTCCCGCGCGGTATTTGGGAGTGTGCTCAAA 3161
DB |||||
DB 3151 GAGCCTACGGCTGGGCTGATCCTGCTTCCCGCGCGGTATTTGGGAGTGTGCTCAAA 3210
QY 3162 TACCGAGAAGTGTGCACGCTGATCCTCTCCACCCACACCTGGATGAGGCAGAGTGTG 3221
DB |||||
DB 3211 TACCGAGAAGTGTGCACGCTGATCCTCTCCACCCACACCTGGATGAGGCAGAGTGTG 3270
QY 3222 GGAGACCGTGTGGCTGTGTGTCAGTGGCGGCTTGTGCTGTGTGGCTCCCACTCTTC 3281
DB |||||
DB 3271 GGAGACCGTGTGGCTGTGTGTCAGTGGCGGCTTGTGCTGTGTGGCTCCCACTCTTC 3330
QY 3282 GTGGCCGTCACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGGCTGCCCTG 3341
DB |||||
DB 3331 CTGGCCGTCACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGGCTGCCCTG 3390
QY 3342 ACCACCAATGAGAAGGCTGACCTGACATGAGGGCAGTGTGGAACACAGGAGGAAAAG 3401
DB |||||
DB 3391 ACCACCAATGAGAAGGCTGACCTGACATGAGGGCAGTGTGGAACACAGGAGGAAAAG 3450
QY 3402 AAGAAATGACAGCGGACAGAGTGTGGCACTCTCAGCTGCTGGCCCTGTATACGAC 3461
DB |||||
DB 3451 AAGAAATGACAGCGGACAGAGTGTGGCACTCTCAGCTGCTGGCCCTGTATACGAC 3510
QY 3462 TGGGTGCCCGGGGCA CGGCTGGTGGAGAGCTGCCACAGAGTGGTGTGCTGTGCC 3521
DB |||||
DB 3511 TGGGTGCCCGGGGCA CGGCTGGTGGAGAGCTGCCACAGAGTGGTGTGCTGTGCC 3570
QY 3522 TACACGGTGGCCATGACGGAGCTTCCGCACTCTTTCCGAGAGCTAGACACGGGCTG 3581
DB |||||
DB 3571 TACACGGTGGCCATGACGGAGCTTCCGCACTCTTTCCGAGAGCTAGACACGGGCTG 3630
QY 3582 GCGAGCTGAGGCTCAGCTGAGTGGGATCTCCGACACAGCCTCGAGGAGTCTTCTG 3641
DB |||||
DB 3631 GCGAGCTGAGGCTCAGCTGAGTGGGATCTCCGACACAGCCTCGAGGAGTCTTCTG 3690
QY 3642 AAGTGTGTGGAGGATGTGCTCGGACACAGATATGGAGGATGGCAGCTGGGGCAGCAC 3701
DB |||||
DB 3691 AAGTGTGTGGAGGATGTGCTCGGACACAGATATGGAGGATGGCAGCTGGGGCAGCAC 3750
QY 3702 CTATGACAGGATGTGCTGGCTAGAGCTAAACCTGTGGGCTCAAGATGCCGCAAGAG 3761
DB |||||
DB 3751 CTATGACAGGATGTGCTGGCTAGAGCTAAACCTGTGGGCTCAAGATGCCGCAAGAG 3810
QY 3762 ACAGGCTGGAGAGGGGACAGCTGGGTGAGCCCGCAGAGACTGACAGGCTCTGGG 3821
DB |||||
DB 3811 ACAGGCTGGAGAGGGGACAGCTGGGTGAGCCCGCAGAGACTGACAGGCTCTGGG 3870
QY 3822 CCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTG 3881
DB |||||
DB 3871 CCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTG 3930
QY 3882 CTTCTAAGCGCTTCTGCTTCCCGCGCAGCGCCCGGCTGTTCGCCCCAGATCGTG 3941
DB |||||
DB 3931 CTTCTAAGCGCTTCTGCTTCCCGCGCAGCGCCCGGCTGTTCGCCCCAGATCGTG 3990
QY 3942 CTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATGCTGCTCTTTTGGG 4001
DB |||||
DB 3991 CTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATGCTGCTCTTTTGGG 4050

QY 4002 CACTACCGGCTCTGGGCTCAGTCCCACTGTA CGGTGCTCAGGTGCTCTTCTTCACT 4061
DB |||||
DB 4051 CACTACCGGCTCTGGGCTCAGTCCCACTGTA CGGTGCTCAGGTGCTCTTCTTCACT 4110
QY 4062 GAGGACGCCCCAGGGGACCTCGAGCTGCCGGCTGTCTGAGGCGCTGCTGAGGAGCA 4121
DB |||||
DB 4111 GAGGACGCCCCAGGGGACCTCGAGCTGCCGGCTGTCTGAGGCGCTGCTGAGGAGCA 4170
QY 4122 GGACTGGAGGAGCCCCCAGTGCAGATAGCTCCCACTGTA CGGTGCTCTGGCACCAAGTTCCT 4181
DB |||||
DB 4171 GGACTGGAGGAGCCCCCAGTGCAGATAGCTCCCACTGTA CGGTGCTCTGGCACCAAGTTCCT 4230
QY 4182 GCTGAAGTGGCAAGGCTCTTGGCCAGTGGCACTGGAACCCAGAGTCTCCATGCCAGCC 4241
DB |||||
DB 4231 GCTGAAGTGGCAAGGCTCTTGGCCAGTGGCACTGGAACCCAGAGTCTCCATGCCAGCC 4290
QY 4242 TGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGTCTGCGGACTGCGCCGCTGCACTGGT 4301
DB |||||
DB 4291 TGCCAGTGTAGCCAGCCCGGTGCCCGGCGCTGTCTGCGGACTGCGCCGCTGCACTGGT 4350
QY 4302 GGTCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGTTCAGAACTCTGACAGGC 4361
DB |||||
DB 4351 GGTCCCTCCCGCCAGGCACTGACCGGCTCTGGGGAAGTGTTCAGAACTCTGACAGGC 4410
QY 4362 CGSAACTGTCTGACTTCTCTGGTCAAGACCTACCGCGCTGTGTGCGCCAGGCGCTGAAG 4421
DB |||||
DB 4411 CGSAACTGTCTGACTTCTCTGGTCAAGACCTACCGCGCTGTGTGCGCCAGGCGCTGAAG 4470
QY 4422 ACTGAAGTGGTGAATGAGTACGGAGCTTCTCGTGGGGGGCGGAGACCA 4481
DB |||||
DB 4471 ACTGAAGTGGTGAATGAGTACGGAGCTTCTCGTGGGGGGCGGAGACCA 4530
QY 4482 GGCCTGCCCTCGGGCAGAGTGTGGCGCTCAGTGGAGGAGTGTGGCGCTGTGAGT 4541
DB |||||
DB 4531 GGCCTGCCCTCGGGCAGAGTGTGGCGCTCAGTGGAGGAGTGTGGCGCTGTGAGT 4590
QY 4542 CCCCTGCCCTGGCGGGCGCTCGACCGTGTCTCTGAAAAA CCTCA CAGCCTGGGCTCACAGC 4601
DB |||||
DB 4591 CCCCTGCCCTGGCGGGCGCTCGACCGTGTCTCTGAAAAA CCTCA CAGCCTGGGCTCACAGC 4650
QY 4602 CTGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAA CAAAGGCTGCACTCATGGTG 4661
DB |||||
DB 4651 CTGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAA CAAAGGCTGCACTCATGGTG 4710
QY 4662 GCCTTTGTCAACCGAGCCAGCAACGCAATCCTCGTGTCTCACCTGCCCGCCAGGCGCGGCC 4721
DB |||||
DB 4711 GCCTTTGTCAACCGAGCCAGCAACGCAATCCTCGTGTCTCACCTGCCCGCCAGGCGCGGCC 4770
QY 4722 CGCCACGCCACAGCATCACCACACTCAAC CACCCCTTGAACTCTCA CCAAGGAGCAGCTG 4781
DB |||||
DB 4771 CGCCACGCCACAGCATCACCACACTCAAC CACCCCTTGAACTCTCA CCAAGGAGCAGCTG 4830
QY 4782 TCTGAGGCTGCACTGATGGCTCTCTGGTGAAGCTCTCGTCTCATCTGTGTGGTCTTT 4841
DB |||||
DB 4831 TCTGAGGCTGCACTGATGGCTCTCTGGTGAAGCTCTCGTCTCATCTGTGTGTGTCTTT 4890
QY 4842 GCAATGCTTTGTCCCGGCCAGCTTCACTCTTCTCTCATTTGAGGAGCAGTCA CCGCA 4901
DB |||||
DB 4891 GCAATGCTTTGTCCCGGCCAGCTTCACTCTTGTCTCATTTGAGGAGCAGTCA CCGCA 4950
QY 4902 GCCAAGCACCTGCACTCATGGGGGCTGTGCCCGCCCTCTACTGTGCTTTGGCAACTTT 4961
DB |||||
DB 4951 GCCAAGCACCTGCACTCATGGGGGCTGTGCCCGCCCTCTACTGTGCTTTGGCAACTTT 5010
QY 4962 CTTGTGGACATGTGTAACTACTTGGTGCAGCATGCACTGCTGTGCTCATCTTTCTGCCC 5021
DB |||||
DB 5011 CTTGTGGACATGTGTAACTACTTGGTGCAGCATGCACTGCTGTGCTCATCTTTCTGCCC 5070
QY 5022 TTCCAGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCTCTGCTGCTGTGCTACTA 5081
DB |||||
DB 5071 TTCCAGCAGAGGCGATATGTGGCCCTGCCAACCTGCTGCTCTCTGCTGCTGTGCTACTA 5130
QY 5082 CTGTATGGCTGGTGCATCACACCGCTCATGTACCCAGCCTCTTCTTCTTCTCGGTGCC 5141

```
Db 5131 CTGTATG----- 5137
|||
Qy 5142 AGCAGCCTATGTGTGCTCACGTGATAAACCTCTTTATTGGCATCAATGGAAGCATG 5201
|||
Db 5138 -GCACAGCCTATGTGTGCTCACGTGATAAACCTCTTTATTGGCATCAATGGAAGCATG 5196
|||
Qy 5202 GCCACCTTTGTGCTTGAGCTCTTCTGTGATCAGAAGCTGCAGGAGGTGAGCCGGATCTTG 5261
|||
Db 5197 GCCACCTTTGTGCTTGAGCTCTTCTGTGATCAGAAGCTGCAGGAGGTGAGCCGGATCTTG 5256
|||
Qy 5262 AAACAGGTCTTCTTATCTTCCCCCACTTCTGCTTGGCGCGGGGCTCATTTGACATGGTG 5321
|||
Db 5257 AAACAGGTCTTCTTATCTTCCCCCACTTCTGCTTGGCGCGGGGCTCATTTGACATGGTG 5316
|||
Qy 5322 CGGAACAGGCCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGAGATTCCAGTCAACC 5381
|||
Db 5317 CGGAACAGGCCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGAGATTCCAGTCAACC 5376
|||
Qy 5382 CTGGCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGATACAGGGGCCCTCTTC 5441
|||
Db 5377 CTGGCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGATACAGGGGCCCTCTTC 5436
|||
Qy 5442 CTTCTCTTCACTACTGTCTGAGCACCGAAAGCCAACTCCTGCCACAGCCAGGGTGAGG 5501
|||
Db 5437 CTTCTCTTCACTACTGTCTGAGCACCGAAAGCCAACTCCTGCCACAGCCAGGGTGAGG 5496
|||
Qy 5502 TCTCTGCCACTCCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGGTGTC 5561
|||
Db 5497 TCTCTGCCACTCCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGGTGTC 5556
|||
Qy 5562 CAAGGAGCCACCCAGGGGATGTGTGGTCTGAGGAACTTGACCAAGGTATACCGTGGG 5621
|||
Db 5557 CAAGGAGCCACCCAGGGGATGTGTGGTCTGAGGAACTTGACCAAGGTATACCGTGGG 5616
|||
Qy 5622 CAGAGGATGCCAGCTGTGTGACCGCTGTGTGCTGGGATTCCTCCCTGGTGAG 5672
|||
Db 5617 CAGAGGATGCCAGCTGTGTGACCGCTGTGTGCTGGGATTCCTCCCTGGTGAG 5667
|||
```

Search completed: December 29, 2004, 23:46:36
Job time : 2933 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 29, 2004, 22:51:51 ; Search time 493 Seconds
(without alignments)
9809.752 Million cell updates/sec

Title: US-09-995-542-4
Perfect score: 6804
Sequence: 1 ctcaggggcgcgctcccc.....cgtgctggtgaaaaa 6804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737.4	25.5	7860	US-09-526-193A-2	Sequence 2, Appli
2	1368	20.1	6819	US-09-032-438C-2	Sequence 2, Appli
3	1368	20.1	7488	US-09-032-438C-1	Sequence 1, Appli
4	1291.4	19.0	6705	US-09-032-438C-5	Sequence 5, Appli
5	920	13.5	8037	US-09-774-528-209	Sequence 209, App
6	404.6	5.9	5894	US-08-665-259-24	Sequence 24, Appli
7	404.6	5.9	5894	US-08-762-500-24	Sequence 24, Appli
8	404.6	5.9	6525	US-08-762-500-74	Sequence 74, Appli
9	161.2	2.4	987	US-09-724-797-7	Sequence 7, Appli
10	158.6	2.3	1062	US-10-000-489-101	Sequence 101, App
11	143.4	2.1	447	US-09-621-976-13889	Sequence 13889, A
12	134.2	2.0	4403765	US-09-103-840A-2	Sequence 2, Appli
13	134.2	2.0	4411529	US-09-103-840A-1	Sequence 1, Appli
14	125.2	1.8	999	US-09-252-991A-4588	Sequence 4588, Ap
15	125.2	1.8	1008	US-09-252-991A-4266	Sequence 4266, Ap
16	120	1.8	36181	US-08-311-731A-120	Sequence 120, App
17	119.8	1.8	20284	US-09-526-193A-21	Sequence 21, Appli
18	119.2	1.8	1767	US-09-489-039A-6408	Sequence 6408, Ap
19	118.4	1.7	1614	US-09-252-991A-6293	Sequence 6293, Ap
20	118.4	1.7	2790	US-09-252-991A-6375	Sequence 6375, Ap
21	118.4	1.7	3948	US-09-252-991A-6175	Sequence 6175, Ap
22	116	1.7	1824	US-09-252-991A-1780	Sequence 1780, Ap
23	113.8	1.7	762	US-09-252-991A-1938	Sequence 1938, Ap
24	110.8	1.6	1614	US-09-252-991A-4198	Sequence 4198, Ap
25	110.2	1.6	4244	US-09-526-193A-27	Sequence 27, Appli
26	107.4	1.6	930	US-09-252-991A-4633	Sequence 4633, Ap
27	107.4	1.6	1179	US-09-252-991A-4241	Sequence 4241, Ap

28	107.4	1.6	1719	4	US-09-252-991A-4534	Sequence 4534, Ap
29	102.8	1.5	2769	4	US-09-489-039A-1767	Sequence 1767, Ap
30	101.2	1.5	6289	4	US-09-526-193A-26	Sequence 26, Appli
31	101.2	1.5	536165	4	US-09-214-808-1	Sequence 1, Appli
32	92	1.4	1176	4	US-09-252-991A-3704	Sequence 3704, Ap
33	92	1.4	2253	4	US-09-252-991A-3810	Sequence 3810, Ap
34	91	1.3	1110	4	US-09-252-991A-8962	Sequence 8962, Ap
35	91	1.3	1278	4	US-09-252-991A-9043	Sequence 9043, Ap
36	91	1.3	1968	4	US-09-252-991A-8743	Sequence 8743, Ap
37	90.4	1.3	750	4	US-09-252-991A-795	Sequence 795, App
38	90.4	1.3	786	4	US-09-252-991A-872	Sequence 872, App
39	90.4	1.3	1197	4	US-09-252-991A-808	Sequence 808, App
40	87.2	1.3	972	4	US-09-489-039A-5325	Sequence 5325, Ap
41	87.2	1.3	1047	4	US-09-602-787A-347	Sequence 347, Appli
42	86.6	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
43	86.6	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
44	85.8	1.3	813	4	US-09-602-787A-349	Sequence 349, App
45	84.6	1.2	35881	4	US-08-311-731A-127	Sequence 127, App

ALIGNMENTS

RESULT 1
US-09-526-193A-2
; Sequence 2, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pimstone, Simon N.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: CHOLESTEROL LEVELS
; FILE REFERENCE: 50110/002005
; CURRENT APPLICATION NUMBER: US/09/526,193A
; CURRENT FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/124,702
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/138,048
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/139,600
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/151,977
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 287
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-526-193A-2

Qy	987	GACAGCAGCTGAGCCCGCTGCTCGAGCTGATTGGAGCCCTGGACAGCCACCCGCTG	1046
Db	1116	GACAACTTCACTCTTACTTCAATGATTGAAGATTTGGAGTTAGTCCTCTT	1175
Qy	1047	TCCGCGCTGCTCTGGAGACGCTGAAGCCTCTGATCTCGGGAAGTACTCTTTGCACCA	1106
Db	1176	TCCGCGATTATCTGGAAGCTCTGAAGCCGCTGCTCGTTGGGAAGATCTGTATACACT	1235
Qy	1107	GATACACTTTTATCCCGAAGCTCATGCCAGGTAAACCGAGCTTCGAGGAGCTACAC	1166
Db	1236	GACACTCCAGCCCAAGGCGAGGTCTGGTGAAGTGAACAGACCTTCAGGAACCTGCT	1295
Qy	1167	CTGCTGAGGATCTCGGGAGGTGTGGAGATGTGGAGCCCGGATCTTACCTTCATG	1226
Db	1296	GTGTTCCATGATCTGGAAGGATGTGGAGGAACCTAGCCCCCAAGACTTGGACCTTCATG	1355
Qy	1227	AACGACAGTTTCAATGTGGCCATGCTGACGGGCTCTCTGCAGATGCAGGA-----	1276

1356 GAGACAGCCAGAAATGGACCTTGTCCGATGCTGTGTGACAGCAGGAGCAATGACCAC 1415
1277 -----TGAAAGGAAGAGGAGGCCAGACCTGGAGGCGCGGACCAACATGG 1320
1416 TTTTGGGAACAGCAGATTTGGATGGCTTAGATTGGACAGCCCAAGACATCGTGGGCTTTTG 1475
1321 ----AGSCCTTGCATCTTCTGGACCCTTGGAGCGGTGGCTACAGCTGGCAGGACGGA 1376
1476 GCCAAGCACCCAGAGGATGCCAGTCCAGTCCAGTAATGGTTCTGTGTACACCTGGAGAGAAGCT 1535
1377 CAGCTGATGTGGGGCACCTGTGGGCAACGCTGGGCGAGTGCAGGAGTCCCTGTCTTG 1436
1536 TTCAAAGAGACTAACAGGCAATCCGAGCAATATCTCGCTTCATGGAGTGTCAACCTG 1595
1437 GA CAAGTGGAGCGGCACCTCTAGAGGACCGCTGGTGTGCGGGCCCTGCAACTGCTC 1496
1596 AACAAAGCTAGAACCCATAGCAACAGAAAGTCTGGCTCATCAACAAAGTCCATGGAGCTGCTG 1655
1497 GCGGAACATCGATTCTGGGCGGGCTGCTCTTCTTGGGACCTGAGGACTCTTCAGACCCC 1556
1656 GATGAGAGGAAGTTCTGGGCTGTAATTGTGTTCACTGGAAATT-----ACTCCAGGCAGC 1709
1557 ACAGAGCA CCAACCCAGACCTGGGCCCGCGGCAACGTGGCATCAAAATCCGATGGAC 1616
1710 ATTGAGC-----TGCCCCATCATGTCAAGTACAGATCCGAATGGAC 1751
1617 ATTGACGTGTGTCAGGAGCAATAGATCAGGGA CAGGTTTTGGGACCTTGCCCCAGGC 1676
1752 ATTGACAAATGGAGAGGACAAATAAAATCAAGATGGGTACTGGGACCCCTGGTCTCGA 1811
1677 GCGGACCCCTGACCGACCTGCGCTAGCTGTGGGGGGCTTGTGTACCTGCAAGACCTG 1736
1812 GCTGACCCCTTTGAGGACATGCGGTAGCTGTGGGGGGCTTGCCTACTTTGCAAGGATGTG 1871
1737 GTGGAGCGTGCAGCGCTGCGGTGCTCAGCGGCGCAACCCCGGGCGGCCTCTACCTG 1796
1872 GTGGAGCAGCAATCATCAGGTGCTGACGGGACCGGACCAAGAAACTGGTGTCTATATG 1931
1797 CAGCAGATGCCCTATCCGTCTATGTGA CAGCGTTCCTGCTGTGTGAGCCGGTGTG 1856
1932 CAACAGATGCCCTATCCCTGTTAGCTTGA TGA CATCTTTCTGGGGGTGATGAGCCGGTCA 1991
1857 CTGCGCTCTTCTGACGCTGGGCTGATCTACTCTCGTGA CACTGACAGTGAAGGCGGTG 1916
1992 ATGCCCTCTTCTATGACGCTGGGCTGGAATTTACTAGTGTCTGTGATCATCAAGGGGCATC 2051
1917 GTGGGAGAGGAGAGCGCGCTGCGGACACCATGCGGCCATGGGGCTCAGCGCGCG 1976
2052 GTGTATGAGAGGAGGACAGGCTGAAGAGACCATGCGGATCATGGGCTGGACACAGC 2111
1977 GTGCTCTGGTAGCTGGTTCTCAGCTGCTCGGGCCCTTCCTGTCTCAGCGCGCGCTG 2036
2112 ATCTCTGGTTTGTAGCTGGTTCA TTAGTGTAGCTCA TTCTCTTCTGTGAGCGCTGGCGCTG 2171
2037 CTGGTCTGTGTCTCAGCTGGGGGACATCTCCCTTACAGCCACCGCGGCTGGTCTTC 2096
2172 CTAGTGTGTCTCTGAAAGTTAGGAAACCTGCTGCTCTACAGTATCCAGCGTGGTGT 2231
2097 CTGTTCTTGGCAGCCTTTCGGGTGGCCACCGTGA CCGAGCTTCCTGTCTCAGCGCTTC 2156
2232 GTCTTCTGTCCGTGTGTGTGTGGTGA CAATCTTCGAGTGTCTCTGATTAGCACTC 2291
2157 TTCTCCCGCCCAACTGGGTGCGGCTGCGGGCGCTGCGGCTACTTCTCCCTCTACCTG 2216
2292 TTCTCAGAGCCAACTGGCAGCAGCCTGTGGGGGACATCATCTACTTCA CGCTGTACCTG 2351
2217 CCCTAGTGTGTGTGTGGCTTGGCGGGA CCGGCTGCCCGGGTGGCCCGCTGGCGCG 2276
2352 CCCTAGTGTGTGTGTGGATGACAGGACTACGTGGGCTTCACTCAAGATCTTGCT 2411
2277 AGCTGTGTGCGCGGTGCTTGGCTTGGCTTGGCTGCGAGAGCCTGGCTGTCTGGAGGAG 2336

2412 AGCTGTGTCTCTCTCTGGCTTTTGGTGTGTGAGTACTTTGCCCCCTTTTGTAGGAG 2471
2337 CAGGGCGAGGGCGCGCAGTGGGCAACAGTGGGCA CCGGGCT---ACGGCAGAGCTTTC 2393
2472 CAGGGCAATTGGAGTGCAGTGGGCAACACTGT TTTGAGAGTCTCTGTGAGGAAGATGGCTTC 2531
2394 AGCTTGCCGCCAGGTCTCTGCGCTTCTGCTGCTGAGCGCGCGCTCTACAGCCCTCGCCACC 2453
2532 AATCTCACCACTTCGTGCTCCATGATGTGTTTGACACCTTCTCTATGGGGTGTATGACC 2591
2454 TGTACTCTGGAAGCTGTGTGCCCAGGCCAGTACGGGATCCCTGAACCATGAAATTTTCT 2513
2592 TGTACATTGAGGCTGTCTTTCAGGCCAGTACGGAAT TCCAGGCCCTCTGTATTTTCT 2651
2514 TTTGCGAGAGACTACTGTGTGGGACCTCGGCCCCCAAGAGTCCAGGCCCTTGCCTCCACC 2573
2652 TGCACCAAGTCTTACTGTGTTTGGCGAGGAAGTGA TGAGAAGAGCACCCCTGGTTTCCAAC 2711
2574 CGCTG---GACCCAAAGGTGCTGTAGAGAGGAC CCGCCGCGCTGTGCTCTGCGTA 2630
2712 CAGAAGAGAAATATCAGAAATCTGCATGGAGAGGAAC CCAACCACTTTGAAGCTGGGCGTG 2771
2631 TCCGTTTCGACAGCTGGAGAAGCGCTTTTCTGGAAG CCGCAGCCAGCCCTTCGCGGGGCTC 2690
2772 TCCATTCAGAACCTGTGTAAGGTCTACCGAGATGGGATGGAAGTGTGTCGATGGCTG 2831
2691 AGCTGTGACTTTCA CAGGGCCACATCA CCGCTTCTCTGGGCCAACAACGGGGCGGCAAG 2750
2832 GCATGAAATTTTATGAGGGCCAGATCACTCTCTCTGCGGCCACAATGGAGCGGGGAAG 2891
2751 ACCACCACTGTCTCATCTTGTAGTGGCTTCTTCC CACCCAGTGGTGGCTCTGCTCTCATC 2810
2892 ACACCACTGATGTAATCTGACCGGGTGT TTTCCCGGACCTCGGGGACCGCTCATC 2951
2811 CTGGGCCACGACGCTCGCTCCAGCATGGCGGC CATCCGGCCCACTGGGCGCTGTGCT 2870
2952 CTGGGAAAGACATTCGCTCTGAGATGAGCACCA TCCGGCAGAACCTGGGGTCTGTCCC 3011
2871 CAGTACAACTGTGTGTGACATGCTGACCGTGGGAC GAGCAGTCTGTGTTCTATGGGCG 2930
3012 CAGCATAACTGTGTGTGACATGCTGACTGTGCTG TCGAAGAACACATCTGTGTTCTATGCGCG 3071
2931 CTGAAGGTCTGAGTGCCTGCTGTGTGGGCGCG CAGCAGCCGCTCTGCTGCAAGATGTG 2990
3072 TTGAAGGCTCTCTGAGAAGCAGTGAAGCGGAGA TGGAGCAGATGGCCCTGGATGTT 3131
2991 GGCTGTGCTCTC---CAAGCAGAGTGTGCAGACT CGCCACCTCTCTGTGGGATGCAACGG 3047
3132 GGTTCGCCATCAAGCAAGCTGAAAGCAAAA CAAGCCAGCTGTCAAGTGGAA TGCAGAGA 3191
3048 AGCTGTCCGTGGCCATTGCTTTGTGGGGGGCT CCAAGTTGTTTATCTTGGAGAGCCT 3107
3192 AAGCTATCTGTGGCTTGGCTTTGTTCGGGGGAT CTAAGGTTGTCTATTCTGATGAACCC 3251
3108 ACGCTGGGGTGTGATCTCTCCCGCGGTA TTTGGGAGCTGTGCTCAATAACCGA 3167
3252 ACGCTGGTGTGACCTTACTCCCGAGGGGAATAT GGGAGCTGTGCTGAAATACCGA 3311
3168 GAAGGTGCACGCTGTATCTCTCCACCCAC CACTGGA TGAGGCAGAGCTGTCTGGGAGAC 3227
3312 CRAAGCGGACCAATATTCTCTACACCA CCAATGGAATGAAAGCGGACGCTCTGGGGAC 3371
3228 CGTGTGGCTGTGGTGGCAGGTGGCCGCTTGTGCT GTGTGGCTTCCCACTCTTCTCTGCGC 3287
3372 AGGATTTGCCATCATCTCTCCATGGGAAGCTGTGCTGTGGGGCTCTCCCTCTTCTGAAG 3431
3288 CGTCACTCGGGCTCGGCTACTACCTGACGCTGTGTAAGSCCGCTGCGCCCTGACCAAC 3347
3432 AACAGCTGGGAACAGGCTACTACCTGACCTTGTGT CAGAAAGATGTGAAATCTCCCTC 3491
3348 AATG-----AGAAAGCTGACACTGACATG 3371
3492 AGTTCTCTGCAGAAACAGTAGTAGCAGTGTGTCTAT CACTGAAAGAGGAGGACAGTGTCT 3551

QY 3372 GAGGCGAGTGTGGACACACAGGACAGAAAGAAAGAAATGGCAGCCAGGGCAGCAGAGTCCGCG 3431
Db |||||
QY 3552 CAGAGCAGTCTCTGATGCTGGCGTGGCGCAGCGACCATGAGAGTGAACACGCTGACCATCGAT 3611
Db |||||
QY 3432 ACTCCTCAGCTGCTGGCCCTGGTACAGCACTGGGTGCGCGGGGACGGCTGGTGGAGAG 3491
Db |||||
QY 3612 GTCTCTGCTATCTCCAACTCATCAAGGAAGCATGTGTCTGAAGCCCGCTGGTGGAGAC 3671
QY 3492 CTGCGACACGAGCTGGTGTGGTGTGCTGCCCTACACGGGTGGCCATGACGGCAGCTTTCGCC 3551
Db |||||
QY 3672 ATAGGGCATGAGCTGACCTATGTGTGCTGCCATATGAAGCTGTAGAGGGAGCCCTTTGTG 3731
QY 3552 ACATCTTCCGAGAGCTAGACACGCGGTGGCGGAGCTGAGGCTCACTGGCTACGGGATC 3611
Db |||||
QY 3732 GAACTCTTTCATGATGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGCGCATC 3791
QY 3612 TCCGACACGAGCTCGAGAGATCTTCTGAAGTGGTGGAGAGTGTGCTGGGACACA 3671
Db |||||
QY 3792 TCAGAGACGAGCCCTGGAAAGAAATATTCCTCAAGGTGGCCGAAGAGTGGGTGGATGCT 3851
QY 3672 GATATGAGGATGGCAGCTGGGGCAGCACCTATGACACAGGCAATTTGCTGGGCTAGACGTA 3731
Db |||||
QY 3852 GAGACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGCGGGCCCTTCGGGACAGCAG 3911
QY 3732 ACCCTGCGGCTCAGATGCCGCCACAGGAGACAGCGCTGGAGAACGG---GGAACACAGCT 3788
Db |||||
QY 3912 AGCTGTCTTCGCCCGCTTCACTGMAAGATGATGCTGCTGATCCAAATGATTTCTGACATAGAC 3971
QY 3789 GGGTCAGCCCCAGAGACTGACAGGGCTCTGGGCCAGACGCGGTGGG-----CGGGTA 3842
Db |||||
QY 3972 CCAGAAATCCAGAGAGACAGACTTGTCTAGTGGATGGATGGCAAGGGTCTCTACAGGTG 4031
QY 3843 CAGGGCTGGGACATGACCCCGCAGCAGCTCCAGGCCCTGCTTCTCAAGGCTTCTGCTT 3902
Db |||||
QY 4032 AAAGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGAAGAGAGACTGCTAAT 4091
QY 3903 GCCCGCCAGCGCGCGCGCTGTTCGCCAGATGCTGCTGCCCTCTTTGTGGGC 3962
Db |||||
QY 4092 GCCAGCGAGTCGGAAGGATTTTGTCTCAGATTGTCTTGCCAGCTGTGTGTCTGCG 4151
QY 3963 CTGGCCCTGCTGTTCAGCTCATGCTGCTCTTTTGGGCACTACCCGGCTCTGGGCTC 4022
Db |||||
QY 4152 ATTGGCTTGTGTTTTCAGCTGATCTGCGCACCTTTGGCAAGTACCCCGAGCTGGAACTT 4211
QY 4023 AGTCCCACTGTACGGTGTCTCAGTGTCTTCTTCACTGAGGAGCGCCCGAGGGACCCCT 4082
Db |||||
QY 4212 CAGCCCTGGATGTACAGAAACAGTACACATTTTGTACGAATGATGCTCTTGAGGACAG 4271
QY 4083 GGAGGTGCCCGCTGCTCAGCGCTGCTCAGGAGGAGGAGTGGAGGAGCCCGCAGTG 4142
Db |||||
QY 4272 GGAACCTTGGAACTCTTAAAGCCCTTACCAAGACCTTGGCTTCGGACCCGCTGTATG 4331
QY 4143 CAGCATAGCTCCCAAGTGTCTCGGCAACAGAAAGTCTCTGTGAAGTGGCCA----- 4194
Db |||||
QY 4332 GAAGGAAACCCAATCCCAAGACAGCCCTCGCAGGCGGGGAGGAAGTGGACCACTGCC 4391
QY 4195 -----AGTCTTGGCCAGTGGCAACTGGACCCCAAGAGTCTTCA 4232
Db |||||
QY 4392 CCAGTTCGCCAGACCATCATGGAACCTCTTCAGAAATGGGAATGGACAATGCAGAACCT 4451
QY 4233 TCCCGAGCTGCGAGTGTAGCGCGCTGCGCGCGCTTGTGCCCGAGTGGCCGGCT 4292
Db |||||
QY 4452 TCACCTGATGCCAGTGTAGCAGGACAAAATCAGAAGATGCTGCTGTGTGCCCA 4511
QY 4293 CGAGTGTGTGTCCTCCCGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTCAGAAC 4352
Db |||||
QY 4512 GGGGCGGGGGCTGCTCTCTCCCAAGAAACAAACAACTGACAGATATCTTCAGGAC 4571
QY 4353 CTGACAGCGCGGAACCTGTCTGACTTCTTGTTCAGACCTACCGCGCTGGTGGCCAG 4412
Db |||||
QY 4572 CTGACAGGAAGAAACATTTCCGATTTATCTGCTGGTGAAGAGTATGTGCAGATCATAGCCAAA 4631

QY 4413 GGCCTGAAGACTAAGAAATGGGTGAATCAGGTACGGTACGGAGGCTTCTCGTGGGGGCG 4472
Db |||||
QY 4632 AGCTTAAAGAACAAAGATCTGGGTGAATGAGTTTAGTATGGCGCTTTTCCCTGGGTGTC 4691
QY 4473 CGAGA---CCAGAGCCTGCCCTCGGCGCAAGAGTTGGGCGGCTCAGTGGAGAGTGTGG 4529
Db |||||
QY 4692 AGTAATATCTCAAGCACCTTCTCCGAGTCAAGAAAGTAAATGATGCCATCAAAACAAATGAAG 4751
QY 4530 GCGTGTCTGAGTCCCTGCTGGGGGGCCCTCGACCGGTGCTCTGAAACACCTCACAGCC 4589
Db |||||
QY 4752 AAAACACCTAAGCTGGCCCAAGACAGTCTTCGAGATCAATTTCTCAACAGCTTGGGAAGA 4811
QY 4590 TGGGCTCACACCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGG 4649
Db |||||
QY 4812 TTTATGACAGGACTGGACACCAAGAAATAATGTCAAGGTGTGTTCAATAACAAGGCTGG 4871
QY 4650 CACTCCATGGTGGCTTTGTCAACGAGCCAGCAACGCAATCTCGGTGCTCACCTGCC 4709
Db |||||
QY 4872 CATGCAATCAGCTCTTTCTGAAATGTTCATCAACAAATGCCATTTCTCCGGGCCAACCTGCAA 4931
QY 4710 CCAGCCCGGCGCCGACGCCACAGCATCAACACACTCAACACACCCCTTGAACTCACCC 4769
Db |||||
QY 4932 AAGGAGAGAACCTTAGCCATTATGGAATTAATGCTTTCAATCATCCCTGAACTCACCC 4991
QY 4770 AAGGAGCAGCTGTCTGAGGCTGCACTGATGCCCTCTCGGTGAGCGTCTCTGTTCCATC 4829
Db |||||
QY 4992 AAGCAGCAGCTCTCAGAGGTGGCTCTGATGACCAATCAGTGGATGTCTTGTGTGCCATC 5051
QY 4830 TGTGTGCTTTTGGCATGTCTTGTCCGGGCCAGCTTCACTCTTGTGCTCTCATTTGAGGAG 4889
Db |||||
QY 5052 TGTGTGCTCTTTGCAATGTCTCTCCAGCCAGCTTTGTGTAATCTCTGATCCAGGAG 5111
QY 4890 CGAGTCAACCCAGCAGCAACGACCTCAGCTCATGGGGGCCCTGTCGCCACCTCTTACTGG 4949
Db |||||
QY 5112 CGGTGACGAAAGCAAAACACTGAGTTTCACTGAGGAGTGAAGCTGTCTACTACTGG 5171
QY 4950 CTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCCAGCATGATCGTGGTGTCTC 5009
Db |||||
QY 5172 CTCTCTAATTTTGTCTGGGATATGTGCAATTTAGCTTGTCTCCCTGCCACACTGCTCATATC 5231
QY 5010 ATCTTCTGGGCTTCCAGCAGAGGCGCATATGTGGCCCTTGGCAACCTGCTGCTCTCTG 5069
Db |||||
QY 5232 ATCTTCTCTGCTTCCAGCAGAGTCTTATGTGCTCTCCACCAATCTGCTGTGTAGCC 5291
QY 5070 CTGTTTGTCTACTGATGTGCTGCTCATCACCGCTCATGTACCCAGCCTCTTCTTCTC 5129
Db |||||
QY 5292 CTCTACTTTTGTGTTATGGGTGTTCAATCAACTCTCATGTACCCAGCCTCTCTTGTG 5351
QY 5130 TTCTCCGTGGCCAGCAGACCTATGTGTGCTCACTGCTCAATAAACCTCTTTATTGGCATC 5189
Db |||||
QY 5352 TTCAAGATCCCGACAGACGCTATGTGTGCTCACAGCGTGAACCTCTTCAATTGGCATT 5411
QY 5190 AATGGAAGCATGGCCACTTTGTGCTTGAAGCTCTTCTCTGATTCAGAACCTGAGGAGTG 5249
Db |||||
QY 5412 AATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTCACCGACAATAAGCTGAATAATC 5471
QY 5250 AGCCGGATCTTGAACAGGCTCTCTTATCTTCCCCCACTTCTGCTTGGCGCGGGGCTC 5309
Db |||||
QY 5472 AATGATATCTTGAAGTCCGTGTCTTGAATCTTCCACATTTTGTGCTGGAGGAGGCTC 5531
QY 5310 ATTGACATGTGCGGAACACAGGCCATGCTGCTTGTAGCGCTTGGGAGACAGGCGAG 5369
Db |||||
QY 5532 ATCGACATGTGAAAACACAGCAATGGCTGATGCCCTGGAAAGGTTTGGGGAGAAATCGC 5591
QY 5370 TTCAGATCAACCCCTGCGTGGGAGGTGCTGGCAAGAAACCTCTTGGCCCATGGTGTATACAG 5429
Db |||||
QY 5592 TTTGTGTCAACCATATCTTGGGAGTGTGGTGGAGCGAAACCTCTTCCGCAATGGCGTGA 5651
QY 5430 GGGCCCTCTTCTTCTTCACTACTGCTGAGCAGCACCGAAGCAACTCTTCCACAG 5489
Db |||||
QY 5652 GGGGTGTGTTTCTTCCCTCATTTACTGTTTTCAGTACAGATTTCTTATCAGGCCCAAGA 5711
QY 5490 CCCAGGTGTAGGTCTCTGCCACTCTCGGGAGAGGAGGAGGATGTAGCCCGGTGAACGG 5549

Qy	1381	CTGATGTGGGGACCTGGTGGGCGACGCTGTGGGCGAGGTGACGAGTGCCTGTCTTGTGGACA	1444
Db	1487	ACTTCGACTGGAGGACATATTTAAACATCACTGATCGCACCTCCGCGCTGGTCAATCAAT	1546
Qy	1441	AGCTGGAGCGGCACCCCTCAGAGCAGCCCTGGTGTGCGGGCCCTGCAACTGCTCGCG	1500
Db	1547	ACCTGGAGTGTGGTCTGTGATGAAGTTTGAAGCTACAATGATGAACCTGAGCTCACCC	1606
Qy	1501	AACATCGATTTCTGGGCGGCGCTCTCTTTCTTGGGACCTGAGGACTCTTTCAGACCCACAG	1560
Db	1607	AACGTGCCCTCTCTACTCTAGGAGAAACATGTTCTGGGCGGAGTGGTATTCCTGACA	1666
Qy	1561	AGCACCAACCCAGACCTGGGCCCCGGCCACGTGGCGCATCAAAATCCGATGGACATTTG	1620
Db	1667	TGTATCCCTGGACACAGCTCTCTACACCCACAGTGAAGTATAAGATCCGAATGGACATAG	1726
Qy	1621	ACGTGCTCACGAGCAACAAATAAGATACGGAGCAGTTTTCGGACCTGCGGCCACGCGCG	1680
Db	1727	ACGTGGTGGAGAAACCAATGAATTTAAGACAGTATTTGGATTTCTGGTCCAGAGCTG	1786
Qy	1681	ACCCCTTGACCGACTCGCTACGTGTGGGGCGCTTCGTGTACTCTGCAAGACTGTGTGG	1740
Db	1787	ATCCCGTGGAGAGTTTCGGGTACATCTGGGGCGGTTTGCTATCTGCAGACATGGTTG	1846
Qy	1741	AGCGTGACCGCTCGCGTGTCTAGCGGCGCAACCCCCGGCGGCGCTCTACTGCGAGC	1800
Db	1847	AACAGGGGATCACAAAGGAGCGAGGTGCGAGCGGAGGCTCCAGTTGGAATCTACTCCAGC	1906
Qy	1801	AGATGCCCTATCCGTGTATGTGACACAGTGTTCCTGCGTGTGCTGAGCGGCTCGTGC	1860
Db	1907	AGATGCCCTATCCCTGTCTCGTGACAGATTTCTTTTCATGATCATCTGTAACCGCTGTTC	1966
Qy	1861	CGCTCTTCTTGACGCTCGCTGGATCTACTCCGTGACACTGACAGTGAAGCGCGTGTGTC	1920
Db	1967	CTATCTTCAATGGTGTCTGGCATGTACTCTGTCTCCATGACTGTGGAAGAGCATGTCT	2026
Qy	1921	GGGAGAGGAGACCGCGCTCGGGACACCAATGCGCGCCATGGGGCTCAGCGCGCGGTGC	1980
Db	2027	TGGAGAGGAGTTTCGAGCTGAAGGAGACCTTGAAATAACAGGGTGTCTCCAATGCAGTGA	2086
Qy	1981	TCTGGCTAGGCTGGTTCTCAGCTGCTCGGCGCCCTTCCTGCTCAGGCGCGCGCTGCTGG	2040
Db	2087	TTTGGTGATCGTGGTTCTTGACAGCTTCTCCATCATGTGCTGATGAGCANTTCTCTCTCTGA	2146
Qy	2041	TTCTGTGTCTCAAGCTGGGGACATCTCTCCCTACAGCCACCGCGGCGTGTCTCTCTGT	2100
Db	2147	CGATATTATCATGCAATGGAAGAACTCTACATTACAGGACCCATTATCTCTTCTCTGT	2206
Qy	2101	TCTTGGCAGCTTCGCGGTGGCCACGGTGACCCAGAGCTTCTGCTCAGCGCCTTCTTCT	2160
Db	2207	TCTTGTGGCTTTCTCCACTGCCACCATCATGTGCTGTTTCTGTCTCAGCACCTTCTTCT	2266
Qy	2161	CCGCGGCAACCTGGCTGCGGCTGCGGCGCTGCGGCTTGGCTTCTCTCTACTGTCCTT	2220
Db	2267	CCAAGGCCAGTCTGGCAGCAGCTGTAGTGGTGTCTATATTTACCCCTCTACTCTGCCAC	2326
Qy	2221	ACGTGCTGTGTGTGGCTTGTGGGACCGGCTGCCCGCGGCTGCGCGTGGCGCGCGAGCC	2280
Db	2327	ACATCTGTGTCTGCTGTGACGACCCGCAATGACCGCTGAGCTGAAGAGGCTGTGAGCT	2386
Qy	2281	TGCTGTGCGCGTGGCTTCGGCTTCGGCTGCGAGAGCCTGGCTCTGCTGGAGGAGCAGG	2340
Db	2387	TACTGTCTCCGGTGGCATTTTGGATTTTGGCACTTGGTGTGCTGTTTGAAGAGCAAG	2446
Qy	2341	GCNAGGCGGCGAGTGGCAAAAGTGGGCAACCGGGCTTACGGCA---GAGCTTTCAGCC	2397
Db	2447	GCCTGGGGCTGCAGTGGAGCAACATCGGGAACAGTCCCAAGGAGGAGCAATTCAGCT	2506
Qy	2398	TGGCCACAGGTCTCTGGCTTCTGTGCTGACGCGGCGCTCTACGGCTCGCCACCTGCT	2457
Db	2507	TCCTGTGTCCATGACAGATGATGCTCTCTGATGCTGTCTATGCTTACTCGCTTGGT	2566
Qy	2458	ACCTGGAAGCTGTGTGCCAGGCGAGTACGGGATCCCTGAACCATGGAAATTTTCTCTTTTC	2517

Db	2557	ACCTTGATCAGGTGTTTCCAGGAGCATATGAAACCCCACTCCTTGGTACTTTCTTCTTAC	2626
Qy	2518	GGAGGAGCTACTGTGGCG-----	2535
Db	2627	AAGAGTCGTATTTGGCTTGGCGGTGAAGGGTGTTCAAACGAGAGAAGAGCCCTTGAAAA	2686
Qy	2536	-GACCTCGGCCCCCAAGAGTCCAGGCCCTTGGCCCCACCCCGCTGGACCCAAAGGTG---	2591
Db	2687	AGACGGAGCCCTTAACGAGGAAACGGAGGATCCAGAGCACCCAGAGAGGTAATACACACT	2746
Qy	2592	--CTGGTAGAAGAGGCAACCGCCCGGCTGAGTCTCTGGCGTATCCGTTGCGAGCTGAGGA	2649
Db	2747	CTTCTTTGAAACGTGAGCATCCAGGGTGGGTCTCTGGGGTATGCGTGAAGAACTCGTAA	2806
Qy	2650	AGCGCTTTCCTGGAAGCCGACGACGCCCTGGGGGCTCAGCCTGGACTTCTTACCAGG	2709
Db	2807	AGATTTTGAAGCCCTGTGGCGGCGACCTGTGAGACCGTCTGAAACATCACCCTTCTACAGA	2866
Qy	2710	GCCACATCACCGCCCTTCTGGGCCAACAAAGGGGCGGCAAGACACCAACCTGTGCATCT	2769
Db	2867	ACCAGATCACCGCATTCCTGGGCCAAATGAGCTGGGAAACCAACCACTTGTCCATCC	2926
Qy	2770	TGAGTGGCCTTTCACCAACCAAGTGTGGCTCTGCTTTCACTCTGGGCCACGAGCTCGCT	2829
Db	2927	TGACGGGTCTGTGGCCACCAACCTCTGGGACTGTGCTCGTTGGGGGAAGGACATTGAAA	2986
Qy	2830	CCAGCATGGCGGCATCCGGCCCCCACTGGGCGTCTGTCTCAGTCAAAACGTGCTGTTG	2889
Db	2987	CCAGCTTGGATGCACTCCGCGAGAGCTTGGCATGTGTCCACAGCAACAATCCTGTGTC	3046
Qy	2890	ACATGCTGACCGTGGACGAGCAGCTCTGTGTTCTATGGGCGCTGAAGGGTCTGAGTGCCG	2949
Db	3047	ACCACCTCAGGTGGCTGAGCAATGCTGTTCTATGCCCAGCTGAAAGGAAGTCCGAGG	3106
Qy	2950	CTGTAGTGGGCCCGAGCAGGACCGTCTGCTGAGGATGTGGGGCTGGTCTCCAAGCAGA	3009
Db	3107	AGGAGGCCAGCTGGAGATGGAAGCCATGTGTGAGGACACAGGCCCTCCACCAAGCGGA	3166
Qy	3010	GTGTGCAGACTGCCCACTCTGTGGGATGCAACGGAAAGCTGTGCGTGCCATGTCCT	3069
Db	3167	ATGAAGAGGCTCAGGACCTATCAGGTGGCATGAGAGAAAGCTGTGCGTTGCCATTGCT	3226
Qy	3070	TTGTGGCGGCTCCCAAGTTGTTATCTCTGACAGAGCTACGGCTGGCGTGGATCTGCTT	3129
Db	3227	TTGTGGGAGATGCCAAGTGTGTGTTCTGACGAAACCCACCTCTGGGGTGGACCTTACT	3286
Qy	3130	CCGCGCCGGTATTTGGGAGCTGCTGCTCAAAATACGAGAAAGTTCGACGCTGATCCTCT	3189
Db	3287	CGAGCGCTCAA TCTGGGATCTGCTCTGAAGTATCGCTCAGGCAGAACCATCATCATGT	3346
Qy	3190	CCACCCACCACTGGATGAGGCAAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTG	3249
Db	3347	CCACTCACCAATGGACGAGCGGACCTCTTGGGGAACCGCATTTGCCATCAITGGCCGAGG	3406
Qy	3250	GCGCTTGTGCTGTGGCTCCCACTCTTCTGCGCGCGTCACTCTGGGCTCCGGCTACT	3309
Db	3407	GAGGCTCTACTGCTCAGGCACCCCACTCTTCTTGAAGAACTGCTTTTGGCAGAGGCTTGT	3466
Qy	3310	ACCTGACGCTGGT-----AAGSCCGCCCTGCCCTTGACCAACCAATGAGAAAGG	3357
Db	3467	ACTTAACTTTGGTCGCAAGATGAAAAACATCAGAGGCCAAAGGAAAGGCAGTGAAGGGA	3526
Qy	3358	CTGACACTGACATGGAGGGCAGTGTGGAACACAGGC-----AGGAAAAGAGA	3405
Db	3527	CTGTGACGCTGCTGCTTAAGGGTTTCTCCACCAACGTTCCAGCCACGTCGATGACCTAA	3586
Qy	3406	ATGGCAGCCAGGCGACGAGAGTCGGCACCTCTCAGCTGTGTCGCCCTGTGACACACTGGG	3465
Db	3587	CTCAGAACAAAGTCTTGGATGGGGATGTAAATGAGCTGATGGATGTAGTGTCTTCCCAATG	3646
Qy	3466	TGCCCGGGGCACGGCTGTGTGGAGAGCTGCCACACAGAGCTGTGTGCTGCTGCCCTACA	3525

Db 3647 TTCCAGAGGCAAAAGCTGGTGGAGTGCAATTGGTCAAGAACTTATCTCTCTTTTCCAAATA 3706
Qy 3526 CGGGTGCCTATGACGGCAGCTTCCGACACTCTTTCGAGAGCTAGACACCGGCTGGCGG 3585
Db 3707 AGAACTTCAAGCAGACAGACATATGCCAGCTTTTTCAGAGAGCTGGAGAGACGCTGGCTG 3766
Qy 3586 AGCTGAGGCTCACTGGCTACGGGATCTCCGACACACAGCTCCGAGGAGATCTTCTGTAAGG 3645
Db 3767 ACTTGGTCTCAGCAGTTTGGAAATTTCTGACACTCCCTGGGAAGATTTTCTGAAGG 3826
Qy 3646 TGGTGGAGGAGTGTGTGCGGACACAGATATGGAGGATGCACTGCGG-----CAGC 3699
Db 3827 TCAGGAGGATTTCTGATTCAGGACCTCTGTTTGGGGTGGCGCTCAGCAGAAAAGAAA 3886
Qy 3700 ACCTATGCA CAGGCAATGTCTGGCTAGACGTAAACCTTGGCGCTCAAGATCGCGCACAGG 3759
Db 3887 ACGTCAACCCCGACACCCCTGCTTGGGTCCAGAGAAAGGCTGGACACACCCCGAGG 3946
Qy 3760 AGACAGGCTGGGAACGGGGAAACAGCTGGGTGACGCCGACAGACTGACACAGGGCTCTG 3819
Db 3947 ACTCAATGTCTGCTCCGAGGGGCGCGGTGCTCACCAGAGGGCAGCTCCGCCAG 4006
Qy 3820 GGCCAGACGCGTGGGCGCGGTATC-----AGGCTGGGCACTGAACCCGACAGCTCC 3873
Db 4007 AGCCAGAGTCCCGAGGCGCGCAGCTCAACACGGGGACACAGCTGGTCTCCAGCATGTGC 4066
Qy 3874 AGGCCGTGCTTCAAGCGCTTCTGCTTGGCGCGCGCAGCGCGCGGCGCTGTTGCGCC 3933
Db 4067 AGGCGCTGCTGCTCAAGAGATTTCCAAACACCACTCCGACGCCCAAGGACTTCTCGCGC 4126
Qy 3934 AGATCGTGTGCTGCGCTCTTTTGTGGCGTGGCGCTGCGCTGCTCAGCTCATCGTGCTC 3993
Db 4127 AGATCGTGTGCTGCGCTTACCTTTGTGTTTTTGTGGCTCTGATGCTTTCTAATTTGTTATCCCTC 4186
Qy 3994 CTTTGGGCACTACCCGCTCTGGGCTCAGTCCCACTATGATAGGTGCTCAGGTGCT 4053
Db 4187 CTTTGGGCAATACCCGCTTTGACCTTCAACCCCTGGATATATGGGACAGTACACCT 4246
Qy 4054 TCTTCAGTAGAGAGCGCCCGAGGAGCCCTGGAGCGTGCGCGGCTGCTCAGGCGGTGCTGC 4113
Db 4247 TCTTCAGCATGGAATGAACACGAGGAGTGAGCAGTTTCAAGTCTGACGACGCTCCTCTGA 4306
Qy 4114 AGGAGGACGACTGGAGGAGCCCGCAGTGACGA----- 4147
Db 4307 ATAAGCCAGGCTTTGGCAACCGCTGCTGAGGAGAGGGTGGCTTCGAGTACCCCTGTG 4366
Qy 4148 -TAGCTCCACAGGTTCTCGGCAACAGAGTTCTGCTGAAGTGGCCAAAGGTCTTGGCCA 4206
Db 4367 GCAACTCAACACCTGGAAGACTCTTCTGTGTCCCAACATCACCCAGCTGTTCCAGA 4426
Qy 4207 GTGGCAACTGGAACCCAGAGTCTCCATCCCGAGCTGCCAGTGTAGCCAGCGGTGCC 4266
Db 4427 AGCAGAAATGACACAGGTCAACCTTCAACATCTGCAAGTGCAGCACCCAGGAGAAC 4486
Qy 4267 GCGGCTGCTGCGGCTGCGGCTGAGTGGTGGTCCCTCCCGCCCGAGGAGTGA 4326
Db 4487 TCACATFCTGCCAGAGTGCCCGAGGGTGGCGGGGCTCCCGCCCGCCCGAGAGAAC 4546
Qy 4327 CCGGCTCTGGGGAAGTGGTTCAAGAACTGACAGGCGGAACTGCTGTGACTTCTGGTCA 4386
Db 4547 AGCGCAGCAGGAAATTTCAAGACCTGACGGGACAGGACATCTCCGACTTCTTGGTAA 4606
Qy 4387 AGACTTACCGCGCTGTGGCGCAGGGCTGGAAGACTAAGAGTGGGTGAATGAGGTCA 4446
Db 4607 AAACGTATCTGCTCTTATAAGAGCAGCTTAAAGAGCAAAATTCGGGTCAATGAACAGA 4666
Qy 4447 GGTACGAGGCTTCTCGTGGGGGCGGAGACCCAGGCGCTGCC---CTCGGGCCAGAGT 4503
Db 4667 GGTATGAGGAATTTCCATTTGAGGAAAGCTCCCGAGTCTGCCCATCAGCGGGAGAC 4726
Qy 4504 TGGGCGGCTCAGTGGAGGAGTTGTGGGCGCTGTGAGTCCCTGCTGCGGGGCGCCCTCG 4563
Db 4727 TTGTTGGGTTTTTAAGCGACCTTGGCCGGATCATGATGTGAGCGGGGCGCTTACTATA 4786

Qy 4564 ACCGTCTCTGTAAGAAACCTCACAGCCTGGGCTCACAGCTGGATGCTCAGGACAGTCTCA 4623
Db 4787 GAGAGCCCTCTAAGAAATACCTGATTTCTTTAAACATCTAGAAATGAGACAACTTA 4846
Qy 4624 AGATCTGGTTTCAACAAAGGCTGGCACTTCAATGGTGGCTTTGTAAACCGAGCCAGCA 4683
Db 4847 AGGTGTGGTTTATAAACAAGGCTGGCATGCGCTGGTCACTTTCTCAATGTGCCACCA 4906
Qy 4684 AGCAATCTCTCGTGTCTACCTGCGCCCGAGGCGCGCCGCGCACAGCCCAAGATCAACA 4743
Db 4907 ACGCCATCTTACGGGCGAGCTGCTTAAGACAGGAGCGCCCGAGGAGTATGGAATCACCG 4966
Qy 4744 CACTCAACACCCCTTGAACCTCACCAAGGAGCAGCTGTCTGAGGCTGCACTCATGAGCCT 4803
Db 4967 TCATTAGCAACCCCTGAACCTGACCAAGAGAGAGCTCTCAGAGATTACAGTCTGACCA 5026
Qy 4804 CTTGGTGGAGCTCTCGTCTCCATCTGTGTGCTTTTGGCCATGTCTTTGTCCCGGCCA 4863
Db 5027 CTTCACTGGATGCTGTGGTTGCCATCTGGCTGATTTTCTCCATGTCTTCTGCTCCAGCA 5086
Qy 4864 GCTTCACTCTTGTCTCTATTTGAGGAGGAGTCAACCGAGCAAGCACTGCGAGCTCATGG 4923
Db 5087 GCTTTGCTCTTTATTTGATCCAGGAGCGGTGAACAAATCCAAAGCACTCCAGTTTATCA 5146
Qy 4924 GGGGCTGTCCCGCCACCTCTACTGCTGTGGCAACTTTCTCTGGGACATGTGTAACACT 4983
Db 5147 GTGGAGTAGCCCCCACCACTACTCTGGTGAACAACTTCTCTGGGACATCATGAATATT 5206
Qy 4984 TGGTGCCAGCATGCTGCTGCTCATCTTTCTGGCCCTTCCAGCAGAGGGCAATGTGG 5043
Db 5207 CCGTAGTGTGGCTGGTGGGCACTTTTCATCGGGTTTCAAGAAAGACCTACACTT 5266
Qy 5044 CCCCTGCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5103
Db 5267 CTCAGAAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5326
Qy 5104 CGCTCATGTACACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5163
Db 5327 CCAATGATGACCCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5386
Qy 5164 CTGCAATAAACCTCTTTTATTTGGCATCAATGGAAGCATGGCCACTTTGTGCTTTGAGTCT 5223
Db 5387 CTTGTCTAATCTGTTTCATCGGCATCAACAGCAGTGTCTATTACTTCTTCTGAAATTA 5446
Qy 5224 T---CTCTGATCAAGAGCTGCGAGGAGTGAAGCGGATCTTTGAAACAGGCTCTCTTAATCT 5280
Db 5447 TTGAGAAATAACCGGACGCTCTCAGGTTCAACCGCGTGTCTGAGGAAAGCTCTCTATTGTCT 5506
Qy 5281 TCCCGCACTTCTGTTGGGCGGGGCTCATTTGACATGTTGGGAAACCGAGCCATGGCTG 5340
Db 5507 TCCCGCACTTCTGCTGGGCGGGGCTCTANTGACCTTGACCTGAGCAGGCTGTGACAG 5566
Qy 5341 ATGCTTTGAGCGCTTTGGGAGACAGGAGTTCAGTCACTCCCTGCGCTGGGAGGTGGTGC 5400
Db 5567 ATGCTATGCCGGTTTGGTGAAGGACCTCTCAAAATCGTTCCACTGGGACCTGATTG 5626
Qy 5401 GCAAGAACTCTTTGGCATGTTGATACAGGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5460
Db 5627 GGAAGAACTGTTTGGCATGTTGTTGGAAGGGGTGGTGTACTTCTCTGACCGCTGCTGG 5686
Qy 5461 TGCAAGCAGGAGCAACTCTCTGCCACAGCCAGGGTGGGCTCTGCGCACTCTCTGGAG 5520
Db 5687 TCAGGCGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5746
Qy 5521 AGGAGGACGAGGATGTAGCCCGTGAACCGGAGCGGGTGGTCCAAAGGAGCAACCGAGGGG 5580
Db 5747 ATGAAGATGATGTGGCTGAAGAAAGACAAAGAAATTAATCTGGTGGAAATAAAACATG 5806
Qy 5581 ATGTGTTGTGCTGAGGAACTTGAACCAAGGTAACCGTGGGCAAGAGATGCCAGCTGTG 5640
Db 5807 ACATCTTAAGGCTTACATGAATAACCAAGATTTATCTGGGCACTCTCCAGGCCAGCTGG 5866

1561 AGCACCCAA CCCAGACCTTGGGCCCGGCCACAGTGGGCATCAAAATCCGATCGAATTG 1620
Db TGATCCCTGGACAGCTCTCTACACCCACAGTGAAGTATAAGATCCGNAATGGACATAG 2314
1621 ACCTGGTTCAGAGGACCAATAAGATCAGGACAGGTTTGGGACCCCTGGCCCGACGCGCG 1680
Db ACCTGGTGGAGAAAACCAATAAGATTAAACACAGGATTGGGATCTGGTCCCAGAGCTG 2374
1681 ACCCTCGACCGACTGGCTGCTAGTGTGGGGGGCTGTGTACCTGCAAGACCTGGTGG 1740
Db ATCCCGTGAAGATTTCCTGGTACATCTGGGGCGGGTTGGCTATCTGACAGACATGGTTG 2434
1741 AGCGTCAGCGCTCCGGTCTCAGCGGCGCCACACCCCGCGCGCTCTACCTCGAGC 1800
Db AACAGGGGATCA CAAGAGACCAAGTCGAGCGGAGCTCAGTTGGAATCTACCTCAGC 2494
1801 AGATGCCCTATCCGTGTATGTGGACAGCTGTCTCGCTGTGCTGAGCCGGTCTGCTGC 1860
Db AGATGCCCTACCCCTGCTCTGTGGACGATTCCTTCATGATCATCTGNAACCGCTGTTCC 2554
1861 CGCTCTTCTGACGCTGGCTGTGATCTACTCCGTGACACTGACAGTGAAGCCGCTGTGC 1920
Db CTATCTTCATGTGTGGCATGGAATCTACTCTGTCTCCATGACTGTGAAGAGCATCGTCT 2614
1921 GGGAGAGGAGACCGGCTGGGGACACCATGCGCGCCATGGGGCTCAGCGCGCGGTGC 1980
Db TGGAGAGGAGTTGCGACTGAAGAGACCTTGAAAAATCAGGGTGTCTCCAAATGCAAGTGA 2674
1981 TCTGGCTAGGCTGTTCTCTCAGCTGCTCGGGGCCCTTCTGCTCAGCGCGCGCTGTGG 2040
Db TTTGGTGTACTGTTCTCTGACAGCTTCTCCATCATGTGATGAGCATCTTCTCTGTA 2734
2041 TTCTGGTCTCAAGCTGGGGGACATCTCTCCCTACAGGACCCCGGGGGTGTCTCTGTT 2100
Db CGATATTCATCATGTCATGGAAGAAATCCTACATTAACAGCGACCCATTCATCTCTCTGT 2794
2101 TCTTGGCAGCTTCGGGTGGCCACCGGTGACCCAGAGCTTCCTGCTCAGCGCGCTCTTCT 2160
Db TCTTGTGGCTTCTCCACTGCCACCATCATGTGTGTCTTCTGCTCAGCACCTTCTTCT 2854
2161 CCGCGCCAACTCGGTGCGGCTGCGGCTGGGCTGCTACTTCTCCCTTACTCGCCCT 2220
Db CCAAGCCAGTCTGGCAGCAGCTGTAGTGTGTCTATTTTCACTCTA CCGCTCTA CCGCCAC 2914
2221 ACCTGTGTGTGTGGCTTGGGGGACCGGCTGCGCGGGGTGGCGGGTGGCGGAGCC 2280
Db ACATCTGTGTCTCGCTGGCAGGACCGCATGACCGCTGAGCTCAAGAGGCTGTGAGCT 2974
2281 TGCTGTGCGCGGTGGCTTGGCTTGGCTGGAGAGCTTGGCTCTGCTGGAGGAGG 2340
Db TACTGTCTCGGTGGCATTTGGAATTTGGCATCTGAGTACCTGGTTCGCTTTGGAAGAGCAAG 3034
2341 GCGAGGCGCGCAGTGGCAACAGTGGGCACCGCGGCTACGGCA---GACGTCTTCAGCC 2397
Db GCCTGGGCTGCACTGGAGCAACATCGGGAACAGTCCCAAGGAGGGGACGAATTCAGCT 3094
2398 TGGCCCAAGTCTCTGGCCCTTCTGCTGTGAACGCGGCTCTACGGGCTCCGCACTGGT 2457
Db TCCTGTCTCATGTCAGATGATGCTCTTGTATGTGCTGTATGGCTTACTGGCTTGGT 3154
2458 ACCTGGAGCTGTGTCAGGCGGCTGACGGATCCCTGAACTGGAATTTTCTTTTC 2517
Db ACCTTGATCAGGTGTTTCCAGGAGACTATGGAACCCCACTTCTTGGTACTTTTCTTCTAC 3214
2518 GGAGGAGCTACTGTGGCG----- 2535
Db AAGAGTGTATTTGGCTTGGCGGTGAAGGTTTCAACAGAGAAAGAGCCCTGGAAA 3274
2536 -GACCTCGGCGCCCAAGAGTCCAGGCCCTTGGCCCAACCCCGCTGGACCCCAAGAGGTG--- 2591
Db AGACCGAGCCCTTAA CAGAGGAAAACGAGGATCCAGAGACCCAGAGGAATATACAGACT 3334

2592 --CTGTTAGAAGAGGACCGCCCGGCTGAGTCTGGCGTATCCGTTTCGAGCCTGGAGA 2649
Db CCTTCTTTGAA CGTGAACATCCAGGGTGGTTCCTGGGGTATCGGTGAAGAACTCTGGTAA 3394
2650 AGCGCTTTCCTGGAAGCCCGCAGCCAGCGCTGCGGGGGCTCAGCCTGGAGCTTCTACAGG 2709
Db AGATTTTTCGAGCCCTTGGCGCGCCAGCTGTGGACCGTCTGAAACATCACCCTCTACGAGA 3454
2710 GGCACATCAACCGCTTCTGGGCGCAACCGGGCCGCAAGACCAACCCCTCTCCATCT 2769
Db ACCAGATACCCGATTTCTGGGCGCAATGGAGCTGGGAAAACCAACACCTTGTTCATCC 3514
2770 TGAGTGGCCTTCTCCACACAGTGTGGTCTGTGCTTCTCTGCGGCCACGACGTCCGCT 2829
Db TGACGGGTCTGTTGCCACCAACCTCTGGGACTGTGCTCGTTGGGGGAAAGGACATTTGAAA 3574
2830 CCAGCATGSCCGCCATTCGGGCCCACTCGGGCGTCTGTCTCAGTACAAAGTGTCTTTTG 2889
Db CCAGCCTGGATGCACTCGGCGAGCGCTTGGCATGTGTCCACAGCAACAACATCTGTTC 3634
2890 ACATGCTGACCTGGACAGCAGCTCTGCTGTATGGCGGCTGAAGGTCTGAGTGCCG 2949
Db ACCACCTCA CGTGGCTGAGCACATGCTGTCTATGCCAGCTGAAAGAAAGTCCOAGG 3694
2950 CTGTAGTGGGCCCGCAGCAGGACCGTCTGTCAGGATGTGGGCTGTGCTCCAAACGAGA 3009
Db AGAGGCCAGCTGGAGATGGAAGCCATGTTGGAGACACAGGCTCCACCACAGGGGA 3754
3010 GTGTGAGACTCGCCACCTCTCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTTGCCCT 3069
Db ATGAAGAGCTCAGGACCTATCAGTGGCATG CAGAGAAAGCTGTCCGTGTTGCCATTTGCCCT 3814
3070 TTGTGGCGGCTCCCAAGTTGTATCTCGAGCAGCTACCGTGGCTGGGATCTGCTT 3129
Db TTGTGGAGATGCAAGGTTGATTTCTGAGCAACCCACCTCTGGGGTGGACCCCTTACT 3874
3130 CCGCGCGGCTATTTGGGAGCTGCTCTCAAAATACGAGAAAGTCCGACGCTGATCTCT 3189
Db CGAGAGCTCAATCTGGGATCTGCTCTGAAATGATCGCTCAGGCAACCAATCATATGT 3934
3190 CCACCCACCACTGTGATGAGGACAGCTCTGGGAGACCGTGTGGCTGTGTGGGAGGTG 3249
Db CCACTCACACATGAGAGGCGGACCTCTTGGGAGCGGCAATGCCATCATTTGCCCGAGG 3994
3250 GCGCTTGTGCTGTGCTGCCCTCCCACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACT 3309
Db GAAGGCTCTACTGCTCAGGACCCCACTCTTCTGAAAGAACTGCTTTGGACAGGCTTGT 4054
3310 ACCTGACGCTGGTG-----AAGGCCCGCTGCGCCCTGACCAACCAATGAGAAGG 3357
Db ACTTAACTTGGTGGCAAGATGAANAACATCCAGAGCAAGAGAAAGCAGTGGGGGA 4114
3358 CTGACACTGACATGAGAGGCGAGTGTGGACACCAAGCG-----AGGAAAAAGAA 3405
Db CTGACGCTGCTGCTTAAGGGTTTCTCCACACGCTGTCCAGCCACGCTCGATGACCTAA 4174
3406 ATGGCAGCCAGGCGACAGAGTCCGACCTCTCAGCTGTGCGCCCTGGTACACACTGGG 3465
Db CTCAGAA CAAGTCTCTGATGGGATGTAAA TGAGCTGTATGATGTAGTTCTCCACATG 4234
3466 TGCCCGGGGCA CGGCTGGTGGAGGAGCTGCAACAGAGCTGTGCTGGTCTGCCCTTACA 3525
Db TTCCAGAGCAAGCTGGTGGAGTGCATTTGGTCAAGAACTTATCTTCTTCTTCCAAATA 4294
3526 CGGGTCCCATGACGCGAGCTTGGCCACACTCTTTCGAGAGCTTAGAACACGCGGCTGGCGG 3585
Db AGAACTTCAAGCACACAGCATATGCCAGCTTTTTCAGAGAGCTGGAGGAGACGCTGGCTG 4354
3586 AGCTGAGGCTCACTGCTACGGGATCTCCGACACAGGACCTCCAGGAGATCTTCTCTGAAG 3645
Db ACCTTGTCTCAGCAGTCTTGGAAATTTCTGACACTCCCTGGAAAGAGATTTTCTGAAGG 4414
3646 TGGTGGAGGAGTGTGCTCGGACACAGATATGGAGGATGGCAGCTGGGG-----CAGC 3699

Qy	5821	ACTGCCCTCAATCCGATGCCAATCTTTGAGCTGTGACGGCCGCGAGCACCTTGGAGCTGC	5880
Db	6635	ACTGTCCTCAGTTTGATGCAATCGATGAGCTGCTACAGGACGAGAACATCTTTACCTTT	6694
Qy	5881	TTGGCGCCTCGCGGTTGCCGGAGCCACAGTTTCCAGACCGCTGGCTCAGGCCTGG	5940
Db	6695	ATGCCCGGCTTCGAGGTTGTACCAAGCAGAAGAAATCGAANAAGTTGCAAACTGGAGTATTA	6754
Qy	5941	CGGCTGTGGACTCTCATGTGTAGCAGACCGGCTCGACGCACCTACAGCGGAGGGAACA	6000
Db	6755	AGAGCCTGGGCTGACTGTCTACGCCGACTGCTGCTGGCTGGCACGTACAGTTGGGGGCAACA	6814
Qy	6001	AACGCAGACTGGCGAAGCGCCCTGGCGCTGTGTGGGACCCAGCCGTGGTGTCTTGGACG	6060
Db	6815	AGCGGAACCTCCACAGCCATGCACTCATTTGGCTGCCACCGCTGGTCTGCTGGATG	6874
Qy	6061	AGCCGACACAGGATGGAACCCAGCGCGGGCGCTTCCTTTGGACAGACCTTTTGGCCG	6120
Db	6875	AGCCACCAAGGATGGACCCCGAGCAACCGCGATGCTGTGGAACGTATCGTGAGCA	6934
Qy	6121	TGGTGGCGGAGGCGGCTTCAGTGATGCTCACTCCCATAGCATGGAGGAGTGTGAAGCGC	6180
Db	6935	TCATCAGAGAAGGAGGCGTGTGGTCTCATATCCACAGCATGGAAGATGTGAGGCAC	6994
Qy	6181	TCTGCTCGCGCTAGCCATCATGGTGAAATGGGCGGTTCCGCTGCTGGGAGACCCCGCAAC	6240
Db	6995	TGTGTACCCGCTGGCCATCATGTTAAAGGCGCGCTTTCGATGTATGGGCACCATTCAGC	7054
Qy	6241	ATCTCAAGGCAATTCGCGCGGTCACACACTGACCTCGGGTGCCCGCGCAGGT	6300
Db	7055	ATCTCAAGTCCAAATTTGGAGATGGCTATATCGTGACAATGAAGATCAAAATCCCCGAAAG	7114
Qy	6301	CCACGCGGAGCGGCGCTT-----CGTGGCGGCGGAGTTCCCTGGGTCGGAGCTGCGCG	6354
Db	7115	ACGACCTGCTTCTTGACTGAACCTGTGGAGAGTTCTTCCAGGGGAATTTCCAGGCA	7174
Qy	6355	AGGCATATGAGGCGCGCTCGCTTCAGCTGCG---CGGGAGGGCGCTGCGCCCTGG	6411
Db	7175	GTGTGCAAGGGGAGAGGCACTACAACATGTCTCCAGTTCCAGGTCCTCTCTCTCCCTCG	7234
Qy	6412	CGGCGTCTTTGGAGAGCTGGCGGTGCAACCGCGCAGAGCACGGCGTGGAGGACTTTTCG	6471
Db	7235	CGAGGATCTCCAGCTCTCTCTCCCAACAGGACGCTGCTCATCGAGGATCTACTAG	7294
Qy	6472	TGAGCCAGACGATGCTGGAGAGGTATTTGTACTTTCTCCAGGACCAAG	6521
Db	7295	TCACACAGACCACTGGACAGGTGTTGTAAATTTTGTCTTAACAGCAG	7344

RESULT 4

RESOL 4
US-09-032-438C-5
; Sequence 5, Application US/09032438C
; Patent No. 6713300
; GENERAL INFORMATION:
; APPLICANT: Rattner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, No. 6713300h F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032.438C

Db 1787 ATCCCGTGAAGATTTCGGGTACATCTGGGGGGGTTTCCCTATCTGCAGGACATGGTTG 1846
Qy 1741 AGCGTGACCGCTCCGCTGCTCAGCGGCCAAACCCCGGGCGCGCTTACCTGCAAGC 1800
Db 1847 AACAGGGATCAAGAGGACAGGTGCAGGCGGAGGCTCCAGTTGGAATCTACCTCCAGC 1906
Qy 1801 AGATGCCCTATCCGCTGATGTGGAGCAGTGTCTCGGTGTGTGAGCCGGTCCGTGC 1860
Db 1907 AGATGCCCTACCCCTGCTTCGTGGAGCAATCTTTCATGATCATCTCCGAACCGCTGTCTCC 1966
Qy 1861 CCTCTTCTGACGCTGGCTCGATCTACTCCGTGACACTGACAGTGAAGGCGGTGTC 1920
Db 1967 CTATCTTCATGTGTGGCATGATCTACTCTGTCTCCATGACTGTGAAGACATCGTCT 2026
Qy 1921 GGGAGAAGAGACGCGGCTGCGGACACCAATCGCGCCATGGGGCTCAGCCGCGGGTGC 1980
Db 2027 TGGAGAAGGAGTTGCGACTGAAGGAGACCTTGAAAAATCAGGGTGTCTCCATGCACTGA 2086
Qy 1981 TCTGGCTAGGCTGGTTCTCAGCTGCTCGGGCCCTTCTGTCTAGCGCGCGGCTGTGG 2040
Db 2087 TTGGTGTACTGGTTCTCTGGACAGCTTCTCCATCATGTCCATGATGACATCTTCTCTCTGA 2146
Qy 2041 TTCTGGTCTCAAGCTGGGGACATCTCCCTACAGCCACCGGGCGTGGTCTTCTGT 2100
Db 2147 CGATATTCATCATGATGAAGAAATCTTACATTACAGGACCCATTCATCTCTCTGT 2206
Qy 2101 TCTTGGCAGCTTCGCGGTGGCCACGGTGCACCCAGAGCTTCTGTCTCAGCGCTTCTTCT 2160
Db 2207 TCTTGTGGCTTCTCCACTGCGCACCATCATGCTGTCTTCTGTCTCAGACCTTCTTCT 2266
Qy 2161 CCGCGGCCAACTGTGCTGGGCTCGGGCGCTGGCCCTTCTTCTTCTTCTTCTTCTTCT 2220
Db 2267 CCAAGGCCAGTCTGCGCAGCAGCTGTAGTGTGTCTCATCTATTCAACCTCTACCTGCCAC 2326
Qy 2221 AGTCTGTGTGTGCTTGGCGGACCGCTGCTCCCGGGTGGCGCGTGGCGCGAGCC 2280
Db 2327 ACATCTGTGTCTCGCTGGCAGACCGCATACCGGTGAGCTGAGTGAAGGCTGTGAGCT 2386
Qy 2281 TGCTGTGCGCGCTTGGCTTGGCTTGGCTGCGAGAGCTGCTGTCTGGAGGAGCAGG 2340
Db 2387 TACTGTCTCCGGTGGCAATTGGATTGGCACTGAGTACTTGGTTCGCTTTGAGAGCAAG 2446
Qy 2341 GCGAGGCGCGAGTGGCAACAGTGGGACCCGCGCTTACGGCA---GACGCTTTCAGCC 2397
Db 2447 GCTCGGGGTGCAGTGGAGCAACATCGGAAACAGTCCACAGGAGGAGCAAAATTCAGCT 2506
Qy 2398 TGGCCAGGTCTCTGGCTTCTGTCTGTGGAGCGCGCTCTACGGCTTCGCCACCTGCT 2457
Db 2507 TCCCTGTCTCCATGACATGATGTCTCTTGAATGCTGTCTATGGCTTACTCGCTTGGT 2566
Qy 2458 ACCTGGAAGCTGTGTGCCAGGCCAGTACGGGATCCCTGMAACATGGAAATTTCTTTTC 2517
Db 2567 ACCTTGATCAGGTGTTTCCAGGAGACTATGNAACCCACTTCTTGTGATCTTCTTCTAC 2626
Qy 2518 GGAGGAGCTA-----CTGGTGC 2534
Db 2627 AAGAGTCGTATTGGCTTGGCGGTGAAGGTGTTCAACACAGAGAAAGAGCCCTGGAAA 2686
Qy 2535 GGAACCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCACCCGCTGGACCCAAAGGTG--- 2591
Db 2687 AGACCGAGCCCTTAAACAGAGGAAACGGAGGATCCAGAGACCCAGAGGAATATACAGACT 2746
Qy 2592 --CTGTGAGAGAGACACCGCGCTGAGTCTCGGGGTATCCGTTCCGAGCCCTGGAGA 2649
Db 2747 CCTTCTTTGAAAGTGAAGCATCAGGGTGGTTCTTGGGGTATGCGTGAAGAAATCTGGTAA 2806
Qy 2650 AGCGCTTTCTTGGAAACCGCGCAGCGCCCTGCGGGGGCTCAGCGTGGACTTCTTACCAAG 2709
Db 2807 AGATTTTGGAGCCGTGCGCGGCCAGCTGTGACCGGTCTGAAATCATTACCTTCTACAGGA 2866
Qy 2710 GCCATCATCGCGCTTCTTGGGCCCAACAGGGGCGGCAAGACCAACCTGTCTCATCT 2769
Db 2867 ACCAGATCACCGCAATCTTGGGCCCAATGGAGCTGGGAAAACCAACCACTTGTCTCATCTCC 2926

Qy 2770 TGAGTGGCTCTTCCACCCAGTGGTGGCTGTGCTTCACTCTGGGCCACGACGTCGCT 2829
Db 2927 TGAGGGTCTGTGGCCAACTCTGGGACTGTGCTCTGTGGGGAGGACATGAAA 2986
Qy 2830 CCAGATGGCGCCATCCGGCCCCACCTGGGGCTGTCTCTCAGTACAACTGTGTTG 2889
Db 2987 CCAGCTGGATGCACTCCGGCAGAGCTTGGCATGTGTCCACAGCAACAATCTCTTCC 3046
Qy 2890 ACATGCTCACCGTGGACAGCAGCTGTGTTCTATGGGGCTGAAGGCTCTGAGTCCG 2949
Db 3047 ACCACTCAGCGTGGCTGAGACATGTGTTCTATGCCAGCTGAAGAAAGTCCCAGG 3106
Qy 2950 CTGTAGTGGGCCCCGAGCAGGACCTGTCTGTCAGGATGTGGGGTGGTCTTCAAGCAGA 3009
Db 3107 AGGAGGCCAGCTGAGATGGAAGCAATGTTGGAGGACACAGGCTCCACCAAGCGGA 3166
Qy 3010 GTGTGACAGCTCGCCACCTCTCTGTGGTGGATGCAACGGAAGCTGTCTGGCCATGTCT 3069
Db 3167 ATGAAGAGGCTCAGGACCTATCAGGTGGCATGACAGAAAGCTGTCTGGTGTGCCCT 3226
Qy 3070 TTGTGGGGGCTCCCAAGTGTATCTCTGGAGGACCTACGGCTGGCGTGGATCTGTCT 3129
Db 3227 TTGTGGGAGATGCCAAGGTGGTGTGATTTGGACGAACCCACCTCTGGGGTGGACCTTACT 3286
Qy 3130 CCGCGCGGTATTTGGGAGCTGTGCTCAAAATACCGAAGGTGCACGCTGTACTCT 3189
Db 3287 CGAGAGCTCAATCTGGGATCTGTCTCTGAGTATCGTCTCAGGCAGAACCATCATCTGT 3346
Qy 3190 COACCCACCACTGGATGAGCAGAGCTGTGGGAGACCGTGTGGCTGTGTGGCAGGTG 3249
Db 3347 CCACTCACCATGGAACGAGGCCGACCTCTTGGGGAACCGCATGCCATCATTTGCCCAGG 3406
Qy 3250 GCGCTTGTGTGTGTGGCTCCCACTCTTCTCTGGCGGCTCACTGGGCTCCGGCTACT 3309
Db 3407 GAAGCTCTACTGTCTCAGGCACCCCACTCTTCTTGAAGAACTGTCTTGGCAGGCTGT 3466
Qy 3310 ACCTGACCTGGTG-----AAGGCGCGCTGCTCCCTGACCACTGAAGAGG 3357
Db 3467 ACTTAACTTGGTGGCAAGATGAATAACATCCAGAGCCAAAGGAAGGAGTGAAGGA 3526
Qy 3358 CTGACACTGACATGAGAGGCGAGTGTGGACACAGG-----CAGGAAAAAGAGA 3405
Db 3527 CCTGAGCTGCTCGTCTAAGGGTTTCTCCACCACTGTCCAGCCCACTGTGATGACTAA 3586
Qy 3406 ATGGCAGCCAGGCGAGAGTCCGCACTCTCAGCTCTGGCCCTGTGACGACTGGG 3465
Db 3587 CTCAGAACAACTCTGGATGGGATGTAAATGAGCTGATGGATGTAGTTCTTCCCACTG 3646
Qy 3466 TGCCCGGGGACCGCTGGTGGAGGAGCTGCCACACGAGCTGTGTGTGCTGCCCTTACA 3525
Db 3647 TTCCAGAGGCAAGCTGGTGGAGTGCAATGGTCAAGAACTTATCTTCTTCTTCCAAATA 3706
Qy 3526 CGGTGCGCATGACGGCAGCTTCCCACTCTTCCGAGAGTACACACGGGCTGGCGG 3585
Db 3707 AGAACTTCAAGCACAGAGCATATGCCAGCTTTTTCAGAGAGCTGGAGAGAGCTGGCTG 3766
Qy 3586 AGCTGAGGCTCATCTGGCTACGGGATCTCCGACACGAGCTCGAGAGATCTTCTGGAAG 3645
Db 3767 ACCTTGTCTCAGCAGTTTTGGAAATTTCTGACACTCCCTCGGAAAGAGATTTTCTGAAG 3826
Qy 3646 TGGTGGAGGAGTGTGTCGGACACAGATATGAGGATGGC-----AGCTGCGGGGAGC 3699
Db 3827 TCACGAGGATTTCTGATTTTCAAGACTCTGTGTTGGGGTGGCGCTCAGCAGAAAGAGAAA 3886
Qy 3700 ACCTATGACAGGCAATGTCTGGCCCTAGACGTAACCTCTGGCTCAAGATGCCGCCACAGG 3759
Db 3887 ACGTCAACCCCGGACACCCCTGTCTGGTCCACAGAGAAAGGCTGGACAGACACCCCAAG 3946
Qy 3760 AGACAGCGCTGGAGAAACGGGAAACAGCTGGGTGAGCCCGCAGAGACTGACAGGGCTCTG 3819
Db 3947 ACTCAATGTCTGTCTCCCGAGGGGCGGGCTGTCTACCCAGAGGGCCAGGCTCTCCCCAG 4006

Db 6080 ACTGCTGGCTGGCAGGTACAGTGGGGGCAACAAGGGAACTCTCCACAGCCATCGCAC 6139
Qy 6028 TGGTTGGGACACAGCGGTGGTGTCTTGGAGAGCGCCACACAGGCATGACCCCGAGG 6087
Db 6140 TCATTGGCTGCCACCGCTGGTGTCTGGATGAGCCACACAGGATGACCCCGAGG 6199
Qy 6088 CGCGGGCTTCTTTGGAAACGCTTTTGGCCGTGGTGGGAGGGCGTTTCAAGTATGC 6147
Db 6200 CACGCCGATGCTGTGGAAACGTCATCGTGAGCATCATCAGAAAGGAGGGGTGTGGTCC 6259
Qy 6148 TCACCTCCCATAGATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGA 6207
Db 6260 TCACATCCACAGCATGGAAGATGTAGGCACTGTGTACCCGGTGGCCATCATGGTAA 6319
Qy 6208 ATGGCGGTTCCGTGCGCTGGGACGCGCAACATCTCAAGGGCAGATTTCGGGGGTC 6267
Db 6320 AGGGCGCTTTTCATGTATGGSCACCATTCAGCATCTCAAGTCCAAATTTGGAGATGCT 6379
Qy 6268 ACACACTGACCTGGGTGGCGCGGCMAGGTCCAGCGGCAGCGGCTTCGTGGCGG 6327
Db 6380 ATATCGTCACAAATGAAGATCAAAATCCCGAAGGACGACCTGTCTTACCTGAACCCCTG 6439
Qy 6328 CGAG-----TTCCCTCGGTGGAGTGGCGAGGSCACATGGAGCGCGCTGCGCTTCC 6381
Db 6440 TGGACAGTTCTTCAGGGGAACTTCCAGGAGTGTGAGAGGAGGACATCAACA 6499
Qy 6382 AGTGCGG---CCGGAGGGCGCTGCGGCCGTGGCGCGGTCTTTTGGAGAGCTGGCGGTGC 6438
Db 6500 TGCTCCAGTTCAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 6559
Qy 6439 ACGGCGCAGAGCAGCGGTGGAGGACTTTTTCGTGAGCGACAGCATGTGAGGAGGTAT 6498
Db 6560 ACAAGGACAGCGCTGCTCATCGAGGAGTACTCAGTCACACAGACCACACTGGACCAGGTGT 6619
Qy 6499 TCTTGACTTCTTCAAGGACGAG 6521
Db 6620 TTGTAAATTTTGTCTAAACAGCAG 6642

RESULT 5

US-09-774-528-209
; Sequence 209, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: Pt_FL_genes Version 2.0
; SEQ ID NO 209
; LENGTH: 8037
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1654)..(7329)
US-09-774-528-209

Query Match 13.5%; Score 920; DB 4; Length 8037;
Best Local Similarity 53.3%; Pred. No. 5,2e-183;
Matches 2832; Conservative 0; Mismatches 2125; Indels 357; Gaps 28;
Qy 1583 CCCGGCCAGTGGGCATCAAAATCCGATGGACATTCAGTGGTGTACAGAGGACCAATAA 1642
Db 1854 CCCGCTCACGTGCACTACAAGATCCGCCAAGACTCCAGTCTTCAACGAGAAACCAACGA 1913
Qy 1643 GATCAGGACAGGTTTTTGGGACCTTGGCCAGCGCGGACCCCTTGCAGCCGCTCGCTA 1702
Db 1914 GATCGCGCGCTTACTGGCGGCTTGGGCCCAATACTGCGCGCGCTTCTACTTCTCTA 1973
Qy 1703 CGTGTGGGGCGGCTTCTGTGTACTCTGCAAGACCTGTGTGGAGGTGCGAGCCGTCGCT 1762
Db 1974 -----CGGCTTGGTCTGGATCCAGGACATGATGGAGCGCGCATCATCGACACTTT 2024
Qy 1763 CAGCGGCGCAACCCCGGG---CGGCTCTACTGTGAGCAGATGCCCTATTCGTGTGTA 1819
Db 2025 TGTGGGCGACACGCTGTGGAGCCAGGCGAGCTAGTGCAGATGTTCCCTTACCCTGCTA 2084
Qy 1820 TGTGAGCAGCGTGTCTGCGTGTGCTGAGCGGCTCGCTCGCGCTTCTTCTGAGCGTGC 1879
Db 2085 CACAGCGCATGACTTCTGTTTGTCAATTGAGCAATGATGCCGCTGTGCATGCTGATCTC 2144
Qy 1880 CTGGATCTACTCCGTGACACTGACAGTGAAAGCGCGTGTGTGGGAGAGAGGAGACGCGCT 1939
Db 2145 CTGGGTCTACTCCGTGGCCATGACCATCCACGACATCGTGGCGGAGAGGAGCACCGGCT 2204
Qy 1940 GCGGACACCATGATCGCGCATGGGGCTCAGCGCGCGGTGTCTGTGGTAGGCTGGTTCT 1999
Db 2205 CAAAGAGGTGATGAAGACCATGGGCTCTGAA CAAAGCGGTGCACTGGGTGGCTGGTTCT 2264
Qy 2000 CAGTGTGCTGGGCGCTTCTGCTCAGCGCGCGCTGCTGTGTTCTGTGCTCAGCTGG 2059
Db 2265 CACCGCTTTGTGAGCTGTCCATCTCGTGACAGCACTCAACGCGATCTCCTGAAGTACGG 2324
Qy 2060 GGACATCTCTCCCTACAGCCACCGCGGCGTGGTCTTCTGTTTCTTGGCAGCGCTTCGCGGT 2119
Db 2325 CCAGTGTCTATGACAGGCAAGTGGTCACTCATCTGGCTCTTCTGGCAGTCTACGCGGT 2384
Qy 2120 GGCCACGGTGACACAGCTTCTGCTCAGCGCTCTTCTTCCCGCGCAACCTGGGTGC 2179
Db 2385 GGCCACCATCATGTTCTGCTTCTGCTGTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2444
Qy 2180 GGCTGCGGGCGCTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2227
Db 2445 GGCTGCGGTGGCATCATCTACTTCTGAGCTAGTGCCTTACATGTACGTGGCGATCCG 2504
Qy 2228 GTGTGTGCTTGGCGGACCGGCTGCGCGGGTGGCGCGCTGGCGGAGCTGCTGCTGC 2287
Db 2505 AGAGAGGTGGCGCATGATAAGATCAAGGCTTCGAGAGTGTGATCGGCTCCTCATGTC 2564
Qy 2288 GCGCGTGGCTTGGCTTGGCTGCGAGAGCTGGCTGTCTGTGGAGGAGGAGGCGAGGG 2347
Db 2565 CACGACGCGCTTGGTCTGGGCTCTAAGTACTTGGCGCTGTATGAGGTGGCGCGGCTGGG 2624
Qy 2348 CGCGCAGTGGCACAACTGGGCA---CCGCGCTTACGCGAGAGCTTTCAGCTGGGCCCA 2404
Db 2625 CATCAGTGGCACACCTTTCAGCCAGTCCCGGTTGGAGGGGAGCGACTTCAACTGTCTCT 2684
Qy 2405 GGTCTCTGGCTTCTGCTGTGAGCAGCGGCGCTCTAGCGGCTCGCCACTGTTACTCTGGA 2464
Db 2685 GGCTGTACCATGTGTGAGTGGAGCGCGTGGTCTATGGCATCTCTCAGTGGTATCTGA 2744
Qy 2465 AGCTGTGCGCCAGGCGAGTACCGGATCCCTGAAACCATGGAATTTTCTTTTTCGAGGAG 2524
Db 2745 GGCTGTGACCCAGGCGATGACGGGCTGCCCGGCTGGTACTTCCCACTGCAGAGATC 2804
Qy 2525 CTACTG-----GTGCGGACCTCGGCCCCCAAGAGTCCAGCGCTTGGCCCCACC 2574
Db 2805 CTACTGGCTGGGCGAGTGGCGGACAGAAAGCTGGGAGTGGAGCTGGCGGTGGCGACGAC 2864

Qy	2575	CGCTGGACCCAAAGGTCGTGG-	2599
Db	2865	CCCCCGCTCAGTGTCTATGGAGGAGGACGAGGCCTGTGTCCATGGAGAGCGGGCTTTTGA	2924
Qy	2596	-----TAGAAGGACACCGCCGGCGCTGAGTCTCTGGCGTATCCGTTCCGAG	2641
Db	2925	GGAGACCCGTGGCATTGAGGAGGAGCCACCCACCTGCCTTGGTTCTTGTCTGTGCGTGGACAA	2984
Qy	2642	CCTGGGAAGCGCTTTCTCTGGAAGCCCGCAGCCAGCCCTCGCGGGGCTCAGCCTTGACACTT	2701
Db	2985	ACTCACCAAGTCTACAGGACGACAGAAGCTGGCCCTGAAACAAAGTGAAGCTTGAACCT	3044
Qy	2702	CTACGAGGCCACATCACCGCTTCTTGGGCCCAACCGGGCCCGCAGAACCAACCAACCTT	2761
Db	3045	CTACGAGAACCAAGTGTGTCTCTTCTTGGGCCCAACCGGGCCGGCAAGACCAACCACT	3104
Qy	2762	GTCATCTTGAATGGCTCTTTCACCAACCAAGTGTGGCTCTGCCCTTCACTCTGCGGCCACGA	2821
Db	3105	GTCATCTCTGACCGGCGCTGTTCCTCTCCAAAGTGGGTTCCGCCCACTATCTACGGGCCACGA	3164
Qy	2822	CGTCGGCTCAGCATGCGCGCATTCGCGGCCCACTCTGGGGTCTGTCTCTCAGTACAACT	2881
Db	3165	CATCCGACCGGATGATGATGATCCGCAAGAACCTGGGATGTGCCCGCAGACATATGT	3224
Qy	2882	GCTGTTTGACATGCTGACCGGTGGACGACGCTCTGTGTTCTATGGGCGGCTGAAGGTCT	2941
Db	3225	GCTCTTTGACCGGCTACCGGTGGAGGAACCTCTGTGTTCTACTCAGGCTCAAGAGCAT	3284
Qy	2942	GAGTGCCTGTAGTGGGCGCGAGCAGGACCGTCTGCTCGAGGATGTGGGCTGTGCTC	3001
Db	3285	GGCTCAGAGGAGATCCGCGAGAGATGGAACAAGATGTCGAGGACCTGGAGCTCTCAA	3344
Qy	3002	CAGCAGAGTGTGCAGACTCGCCACCTCTGTGGGATGCAACGGAAGCTGTCCGTTGC	3061
Db	3345	CAAAAGCACTCACTGTGTGAGACATTGTGGGGTGGGATGAAGCGCAAGCTGTCCGTTGC	3404
Qy	3062	CATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTGGACGAGCTACGCTGGCGTGA	3121
Db	3405	CATCGCTTCTGTGGCGGCTCTCGCGCCATCATCTGGAAGCCACGCGGGCGTGA	3464
Qy	3122	TCTGTCTTCCGCGCGGATTTGGGAGCTGTGCTCAATACCGAAGAGGTGCAAGCT	3181
Db	3465	CCCCTAGCGCGCGCCATCTGGGACCTCATCTGAAGTACAAGCCAGCGCGCACCAT	3524
Qy	3182	GATCCTCTCCACCCACCACTGGATGAGGACAGCTGCTCGGAGACCGTGTGCTGTGCT	3241
Db	3525	CTTCTCTGCACCCACACATGGATGAGGTGACTGCTTGGGACCGCATTGCATCAT	3584
Qy	3242	GGCAGGTGGCGCTGTGTCTGTGCTGTGCTCTCCCACTCTCTGCGCGCTCACCTGGGCTC	3301
Db	3585	CTCCATGGGAAGCTCAAGTGTGCGGCTCCCGCTCTTCTCAAGGGCACCTATGGCGA	3644
Qy	3302	CGGCTACTACTGACGCTGTGGAAGCCCGCTGCCCCCTGACCACTTAATGAGAAGGTGA	3361
Db	3645	CGGGTACCGCTCACTGCTGTGTCGAAGCGCGCCGCGAGCGGGGGCCCCCAAGAGCCAGG	3704
Qy	3362	CACTGAATGGAGGGCAGTGTGGACACAGGCAGGAAAAGAAATGGCAGCCAGGGCAG	3421
Db	3705	GCTGGCATCCAGCCCCCAGGTGGGCGCCCGCTGAGCAGTGTCTCGAGCTCCAGGTGTC	3764
Qy	3422	CAGAGTCGGCACTCCTCAGCTGTCTGGCCCTGTGTACAGCACTGGGTGCCCCGGGCACT	3481
Db	3765	CCAGTTTATCCGAAGCATGTGGCTCTTGCTGCTGCTCTCAGACACAGCACTGAGCT	3824
Qy	3482	GGTGGAGGAGCTGCCACACGAGCTGGTG-----CTGTGTGTGCC	3520
Db	3825	CTCCTATATCTGCCCCAGGCGCGCCCAAGAGGGGGCTTTTCGAGCGCTCTTTCAGCA	3884
Qy	3521	CTACACGGTGCCTATGCGCAGCTTTGCCACACTCTTCCGAGAGCTAGACACGCGGCT	3580
Db	3885	CCTGGAGCGAGCTGTGATGCACTGCACCTCAGACGCTTCTGGGCTGTATGACACGACCT	3944
Qy	3581	GGCGGA-----GCTGAGGCTCACTGGCTACGG	3607

Db 5025 CGGCGACAAATGCTCTGAGTACCTGCTCTTACCTCCGACCGCTTCCAGTGCACCGGTA 5084
Qy 4551 GCGCG- - - - - GCGCCTCGA CCGGTGCTCTGAAACCTCTACAGCCTGGGCTCA CAGCCTGGA 4606
Db 5085 TGGGGCCATACCTTTTGGAAACGCTCTGAAGTCCATCCAGCCTCATTTGGACACCGGCG 5144
Qy 4607 TGCTCAGACAGTCTCAAGATCTGTT- - - - - 4642
Db 5145 CCCACCCATGGTCGCGAAGATCGCGGTGCGAGGGCTGCCAGGTTTCTTACAAACAA 5204
Qy 4643 AGGCTGGCACTCCATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCTCGTGTCA 4702
Db 5205 GGGCTATCAGAGATGCCCCACTTACTCAACAGCCTCAACAGCCATCCCTGCGTGGCCAA 5264
Qy 4703 CTTGCCCCAGCGCCGCGCCGCG- - - - - ACGCCCAACAGCATCAACACATCAACACCCCTT 4759
Db 5265 CTTGCCCAAGAGAGAGGCGCAACCCGCGGCTTACCGCATCA CCGTCAACACACCCCAT 5324
Qy 4760 GAACTCTCAACGAAGAGAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGGTGGAGTCC 4819
Db 5325 GAATAAGACCAAGCGCCAGCCTCTCCTGGAATTACCTGCTGAGGGCAC- - - - - GATGTCGT 5381
Qy 4820 CGTCTCCATCTGTGTGGTCTTTGCGCATCTCTTTGCTCCGCGCCAGCTTCACTCTGTCT 4879
Db 5382 CATGCCCATCTTCATCATCTGTGGCCATCTCTCTGTCGCGCCAGCTTCTGTCTTCC 5441
Qy 4880 CATTGAGAGCGAGTCAACCGAGCCAAAGCACTGAGCTCATGCGGGGCTGTCCCCAC 4939
Db 5442 CGTGGCCAGAGTCCAACAGGCCAAGCATCTGCAAGTTGTGACGCGCTGCAACCCCAT 5501
Qy 4940 CCTTACTGGCTTGGCAACTTTCTCTGGGACAT- - - - - GTGTAACTA 4981
Db 5502 CATCTACTGGCTGGCGAACTAGTGTGGGACATGCGCGCCCAACCCCTGCGAGCTCAA 5561
Qy 4982 CTTGGTGGCAGATGATCGTGTGCTCATCTTCTGSCCTTCCAGCAGAGGGCATATGT 5041
Db 5562 CTTGGTCCCCGCTACCTCTGTGTGCATCATCTGTGTGCACTGCGCGGCTACAC 5621
Qy 5042 GCGCCCTGCCAACCTGCTGCTCTCTGCTGTGTACTACTGTATGGCTGGTGCATCAC 5101
Db 5622 GTGCGCCACCAACTTCCCTGCGCTCTCTCTCTCTCTGCTCTATGGTGGTCCATCAC 5681
Qy 5102 ACCGCTCATGACCCAGCCTCTTCTCTTCTTCGGTGCCAGCAGCCTATGTGGTCT 5161
Db 5682 GCCCATCATGTACCCGGCTCTCTCTGCTGTGAGGTGCCAGCTCCGCTACGTGTCTCT 5741
Qy 5162 CACCTGCAATAACCTCTTTATGGAATCAATGGAAGATGCGACCTTTGTGCTTGAGCT 5221
Db 5742 CATTTGTCATCAATCTCTTTCATCGGCATCAACCGCCACCGTGGCCACCTTCTCTGCTACAGCT 5801
Qy 5222 CTTC- - - - - TCTGATCAGAACTGACGAGGTGAGCCGGATCTTGAAACAGGTCTTCCCTTAT 5278
Db 5802 CTTTGAGCAGCAAGGACCTGAAAGTTGTCAACAGTTACTTGAAAGCTGCTTCTCTCAT 5861
Qy 5279 CTTCCCCCATCTCTCTTGGCCGCGGGCTCAITGACATGGTGGGAAACCAAGGCCATGGC 5338
Db 5862 TTTCCCCCAACTACAACCTGGGCCACCGGCTCATGAGATGSCCTTACAAGAGTACATCA 5921
Qy 5339 TGATGCTTTGAGCGCTTGGGAGACAGGC- - - - - AGTTCAGTCAACCCCTGCGCTGGAGGT 5395
Db 5922 CGAGTACTACGCCAAGATTGGCCAGTTTGAACAAGTGAAGTCCCGCTTTCGAGTGGGACAT 5981
Qy 5396 GGTGGCAAGAACCTCTTGGCCATGGTATACAGGGGCCCTCTCTCTCTCTTCTTCACT 5455
Db 5982 TGTCAACCGCGACTGTGGCCATGGCGGTGTGAGGGCGCTGTGGGTTTCTCTCTGACCAT 6041
Qy 5456 ACTGTGACGACCGGAACCAACTCTTCCCAACAGCCAGGCTGAGGTCTCTGCCACTCT 5515
Db 6042 CATGTGCCAGTACAATCTCTCGCGCGCCACAGGCATGCTGTGTCT- - - - - ACCAAGCC 6098
Qy 5516 GGAAGAGAGCAAGAGATGTAGCCGTGAACGGAGCGGTGTGCCAAGAGCAACCCA 5575
Db 6099 TGTGGAGGATGTGGACGTGGCGCAGTGAAGCGGCGAGGTGCTCCGGGAGACGCCGA 6158

Qy 5576 GGGGATGTGTTGTTGCTGAGAACTTGACCAAGTATACCGTGGGCAGA- - - - - G 5626
Db 6159 CAATGACATGCTCAAGATTGAGAACTTGACCAAGTCTTACAAGTCCCGAAGATTGGCG 6218
Qy 5627 GATGCGAGCTTTGACCGCTTGTGCTGGGATTTCCCTCTGTTGAGTGTTTTGGCTGCT 5686
Db 6219 TATCTGCGCTTTGACCGCTTGTGCTGGTGTGCTCTGCGAGTCTTTCGGGCTCT 6278
Qy 5687 GGGTGTGAATGGACAGGAGAGCTCCACGTTTTCGATGTTGACGGGGACACATTGGC 5746
Db 6279 GGGCGTCAACGCTGGGCAAGACCAAGACCTTCAAGATGCTGACCGCGCAGAGACAC 6338
Qy 5747 CAGCAGGCGCGAGGCTGTGCTGGCAGGCCACAGCGGTGGCCCGGAAACCCAGTGTCTGCCA 5806
Db 6339 GACGGGGCGGAGGCTTTCGTCATGAGACACAGCGTGTGAAAGAGCTGTCTCCAGTGA 6398
Qy 5807 CTTCAGCATGGGATGACTGCTCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGA 5866
Db 6399 GCAGAGCCTCGGCTACTGCGCGCAGTGTGACGCGCTGTTTCGACGAGCTCACGGGCCGA 6458
Qy 5867 GCACCTGAGCTGCTTGGCGCCTGCGGCTGTCCCGAGGCCAGGTTGCCAGACCGC 5926
Db 6459 GCACCTGAGCTGTACACGCGCTGCGTGGGATCTCTTGAAGGACGAGGCCGCGTGGT 6518
Qy 5927 TGGTCTCAGGCTGCGGCTCTGGGACTCTCATGTTAGCAGACCGGCTGCGAGGCACCTA 5986
Db 6519 GAAGTGGGCTCTGAGGAAGCTGGAGCTGACCAAGTACGACAGCAAGCGGCTGGCACCTA 6578
Qy 5987 CAGCGAGGGAAACAAACGCAAGCTGGGCAACGCGCTGCGCTGTTGGGACCCAGCCGT 6046
Db 6579 CAGCGCGGCAACAAAGGGAAGCTCTCAGCGCCATCGCCCTCATTTGGGTACCCAGCCTT 6638
Qy 6047 GGTGTTTCTGACAGCGGACCAAGGATGGAACCGAGCGCGGCGCTTCTTTGAA 6106
Db 6639 CATCTCTGAGCAGGCCACCAAGGATGGAACCGGCGCTTCTCTCTGAA 6698
Qy 6107 CAGCCTTTTGGCCGTGTTGGGAGGCGGTTTCACTGATGCTCACTCCCATAGCATGGA 6166
Db 6699 CCTCATCTCGACCTCATCAAGACAGGCGGTTCACTGGTGTGATCATCACAGCATGGA 6758
Qy 6167 GGAAGTGAAGCGCTCTGCTGCGCTTAGCATCATGTTGAATGGGGTTCGCTGCT 6226
Db 6759 GGAAGTGAAGCGCTGTGACGCGCTGGCCATCATGTTGAACGCTGCTGCGTGCCT 6818
Qy 6227 GGGCAGCCCGCAACATCTCAAGGCGAGTTCCGCGCGGCTCACACTGACCTTCCG- - - 6284
Db 6819 GGGCAGCATCCAGCACTGAAGAACCGGTTTGGAGATGGCTTACATGATCAACGCTGCGAC 6878
Qy 6285 -GTCCCGCGCAAGTCCAGCGCGGCGCTTCTGTCGCGCGGCTTCCCTGGGTC 6343
Db 6879 CAAGACGACGAGTGTGAAGACGCTGCTGCTGCTTCTTCAACCGCAACTTCCCGAAGC 6938
Qy 6344 GGAAGTGGGAGCAGCATGAGGCGGCTTCCGCTTCCAGCTGCGCGGAGGCGCTG 6403
Db 6939 CATGCTCAAGAGCGGACCAACAAAGGTGCACTACAGCTCAAGTCG- - - - - GAGCAT 6995
Qy 6404 CGCCTGGCGGCGCTTTTGGAGAGCTGGCGGTGCAAGCGCGCAGACAGCGCGTGGAGGA 6463
Db 6996 CTGCTGGCGGCGGCTTTCAGCAAGATGAGAGAGGTGCTGCTGGGCTGCTGCGCATCGAGGA 7055
Qy 6464 CTTTTCGCTGAGCAGACGATGCTGGAGGAGTATCTTGTACTTCTTCAAGGACCGAGG 6523
Db 7056 CTACTCGCTCAGCCAGACCACTGAGCAATGTGTTCTGTGAACTTTGCCAAGAACAGAG 7115
Qy 6524 GAAGGACGAGACACCGAAGAGCAGAGAGGCA 6557
Db 7116 TGACAACTTGAGCAGCAGGAGAGCGGACCGCCA 7149

RESULT 6
US-08-665-259-24
; Sequence 24, Application US/08665259

Db	1744	GTGTTACAGTCCGGAAGAGCCCTGGCGCCTGTGCCCCGACGACGACATCCTGTTTGACAACTT	1803
Qy	2897	GACCGTGGACGAGCACGCTCTGTTTCTATATGGCGGCTGAAGGCTCTAGTGCCGCTGTAGT	2956
Db	1804	GACAGTCGCAAGACACCTTTATTTTCTACGCCCAAGCTGAAGGGCCTGTCACTCAGAAAGTG	1863
Qy	2957	GGGCCCCGACGACGAGACCGTCTGCTGTCAGGATGTGTGGGCTGGTCTTCCAAAGCAGAGTGTGCA	3016
Db	1864	CCCTGAAGAAGTCAAGCAGATGCTGCGCATCATCGGCCCTGGAGCAAGTGGAACTCACC	1923
Qy	3017	GACTTCGCCACCTCTCTCGTGGGATGCAACGGAAGCTCTCGGTGGCCCATTCGCTTTGTGGG	3076
Db	1924	GAGCGCTTCTCTGAGCGGGGCATGAGCGCAAGCTCTCCATCGGCATCGCCCTCATCGC	1983
Qy	3077	CGGCTCCCAAGTTGTTATCTTGAAGCAGAGCTTACGGCTGGCGTGGATCCTGTTCCCGCGG	3136
Db	1984	AGGCTCCAAGTGCTGTATGTGAGCAGAGCCCACTCGGGCATGGACGCCATCTCCAGGAG	2043
Qy	3137	CGGTATTTGGGAGCTGTGCTCAATACCAGGAAGTTCGACGCTGATCTCTCCACCCA	3196
Db	2044	GGCCATCTGGGATCTTTCTTACGGCGCAGAAAAGTGAACGGACCAATCGTGTGACCAACCA	2103
Qy	3197	CCACCTGGATGAGCAGAGCTGTCTGGAGACCGGTGTGGCTGTGGTGTGGAGTGGCCGCTT	3256
Db	2104	CTTCAATGGACGAGGCTGACCTGCTTGGGAGACCGCATCGCCATCATGTCGCCAAGGGGGAGCT	2163
Qy	3257	GTGCTGTGTGGCTCCCACTCTTTCTTCCGCGCGCTCACCTGGGCTCCGGCTACTACCTGAC	3316
Db	2164	GCAGTGTCTGGGTCTCTCGCTGTTTCTTCAAGCAGAAATACGGTCCCGGCTATCAATGAC	2223
Qy	3317	GCTGGTGAAGG	3327
Db	2224	GCTGGTGAAGG	2234

RESULT 8

US-08-762-500-74
, Sequence 74, Application US/08762500
, Patent No. 6030806
, GENERAL INFORMATION:
, APPLICANT: Landes, Gregory M.
, APPLICANT: Burn, Timothy C.
, APPLICANT: Connors, Timothy D.
, APPLICANT: Dackowski, William R.
, APPLICANT: Van Raay, Katherine J.
, APPLICANT: Klinger, Katherine W.
, TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
, TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
, NUMBER OF SEQUENCES: 83
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: GENZYME CORPORATION
, STREET: One Mountain Road
, CITY: Framingham
, STATE: Massachusetts
, COUNTRY: United States of America
, ZIP: 01701
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/762,500
, FILING DATE: 09-DEC-1996
, CLASSIFICATION: 435
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/665,259
, FILING DATE: 17-JUN-1996
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: PCT/US96/10469
, FILING DATE: 17-JUN-1996
, ATTORNEY/AGENT INFORMATION:
, NAME: Dugan, Deborah A.

QY 2549 -----CAAGAGTCAGCCCTTGGCCCAACCCCGCTGGACCCCAAGGTGCTGGTAGAAGA 2602
Db AGGGAAGAGGAGGAAGAGCAGTGACCCCGAGAGCACTCAGAAACAGTAGTACTTTGAAGC 2134
QY 2603 GGCACCGCCCGCCTGAGTCTCTGGCGTATCCGTTCCGACGCTGGAGAGCCCTTTCCTGG 2662
Db CGAGCCAGAGACCTTGGTGGCGGGATCAAGATCAAGCACCTGTCCAAGGTGTTCAAGGT 2194
QY 2663 AAGCCCGCA-----GCCAGCCCTCGCGGGCTCAGCCTGGACTTCTACAGGGGCCACAT 2716
Db GGGAAATGAAGACAGGGCGCGCTCAGAGACTGAACTCAACCTCAACCTGTACGAGGACAGAT 2254
QY 2717 CACCGCCCTTCTGGGCCCAACACGGGGCGGCAAGACCAACCCCTGTCCATCTTGAAGTGG 2776
Db CACCGCTCTGTGGGCCCAACACGGGTGCGGGGAAGACCAACCCCTCTCATGTCTACAGG 2314
QY 2777 CTTCTTCCACCCAGTGTGCTCTGCTTCTATCTGGGCCACGACGCTCCGCTCCAGCAT 2836
Db TCTCTTTTCCCCCACCAGTGAACGGGCATACATCAGCGGGTATGAAATTTTCCAGGACAT 2374
QY 2837 GCGCGCCATCGCGCCCACTCGGGGCTGTCTCTCAGTACAAACGCTGCTGTTTGACATGCT 2896
Db GGTTCAGATCCGAGAGCTGGGCTGTGCCCGCAGCAGCATCTCTGTTTGACAACTT 2434
QY 2897 GACGTGAGCAGACAGTCTGGTTCTATGGCGGCTGAAGGGTCTGAGTGCCTGTAGT 2956
Db GACAGTCGACAGCACCTTTATTTCTAGCCCGCAGCTGAAGGGCCTGTCACTCAGAGAGTG 2494
QY 2957 GGGCCCCCAGCAGGACCTCTGCTGCAGAGATGTGGGGCTGTCTTCAAGCAGAGTGTGCA 3016
Db CCCTGAAGAAAGTCAAGCAGATGCTGCACATCATCGGCCCTGGAGGACAAAGTGGAAGTCA 2554
QY 3017 GACTCGCCACCTCTCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTTGCTTTGGGG 3076
Db GAGCGCTTCTTACGCGGGGATGAGCGGCAAGCTCTCCATCGGCATCGCCCTCATCGC 2614
QY 3077 CGGCTCCCAAGTGTATCTTGGAGCAGCTACGGCTGGCGTGGATCTGCTTCCCGCG 3136
Db AGGCTCCAAGTGTCTGACTTGGAGAGCCACCTCGGGCATGAGCGCATCTCCAGGAG 2674
QY 3137 CGGTATTTGGAGCTGTCTCAATACCGAGAGGTGCGAGGTGATCTCTCCAGCCA 3196
Db GGCATCTGGGATCTTCTTACGCGCGAAGAAAGTGACCGCACCATCTGCTGACCAACCCA 2734
QY 3197 CCACCTGATCAGGCAGAGCTGTGGGAGACCGTGTGCTGTGGTGGCAGGTGCCGCTT 3256
Db CTTATGAGCAGGCTGACCTGTGGGAGACCGCATCGCCATCATGGCCAAAGGGGAGCT 2794
QY 3257 GTGCTGTGTGGCTCCCACTCTTCTGCGCGCTACCTGGGCTCCGGCTACTACTGAC 3316
Db GCAGTGTGCGGGTCTCGCTTCTTCAAGCAGAAATACGGTGCAGGCTATCATGAC 2854
QY 3317 GCTGGTGAAG 3327
Db GCTGGTGAAG 2865

RESULT 9
US-09-724-797-7
; Sequence 7, Application US/09724797
; Patent No. 673398
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(987)
US-09-724-797-7

Query Match 2.4%; Score 161.2; DB 4; Length 987;
Best Local Similarity 50.5%; Pred. No. 1.6e-24; Indels 12; Gaps 3;
Matches 475; Conservative 0; Mismatches 453;

QY 5627 GATGCCAGCTTGTGACCGCTTGTGCTGGGATTTCCCTCTGCTGAGTGTGTTGGGCTGCT 5686
Db 60 GACGGCGCGGTTCAGCGACTGGACCTGGTGGTGGGGCGGGACGATCCACGGGGTGCT 119
QY 5687 GGGTGTGAATGGAGCAGGGAAGAGCTCCACGTTTTCGATGTGACGGGGGACACATTTGGC 5746
Db 120 GGGCCCAAGCGGCCCGGCAAGACGACGGCCATCAAGATGCTCGCCACGCTGATGCCGACC 179
QY 5747 CAGCAGGGGGAGGCTGTGCTGGCAGGCCACAGCGTGGCCGGGAAACCCAGTGTGCGCA 5806
Db 180 CACCTCCGCGCACCCGCTCCGTGCTGGGGCAGCAGTGTCCCGAGGCCCGGAGGTCCG 239
QY 5807 CCTCAGCATGGATPACTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGCGA 5866
Db 240 GCGCGCATCGGCTCACCGGCCAGACCATGTCCGTGACGAGGACATGACCGGGGTGCA 299
QY 5867 GCACCTGAGAGCTGTTGGGGCTGTGGCGGTGTCCTGGAGGCCAGGTGTCAGACCGC 5926
Db 300 GAACTTATCTCCCGCGCGCTTGCAGGGTCTCGGCAACGCTCCGCGGCCGCGCGGC 359
QY 5927 TGGCTCAGGCTGCGGCTGTGGGACTCTCATGTATGTCAGACCGGCTGCGAGCACCTA 5986
Db 360 GGAGCAGTTGATGGAGGCTTCGACCTCACCGAGGTGCGCGCGCGCTGTGTGAAGACCTT 419
QY 5987 CAGCGGAGGGAACAAACGCAAGCTTGGGCAACCGCCCTGCGCTGTGGGACCCAGACCGT 6046
Db 420 CTCGGCGGGCAGCGGCGGCGCATCGACGTGGCGCGGAGCATGTGTGTCAACCCCGAGCT 479
QY 6047 GGTGTTTCTGGAACGAGCGGACCAAGGATGGAACCCAGAGCGCGGGCTTCTTTGAA 6106
Db 480 GCTGTTCTCGACAGCGGACCAACCGGCTCGACCCCGCAGCGCGGACGCGAGGTCTGGGA 539
QY 6107 CAGCTTTTGGCCGTGCTGGGAGGGCGCTTCACTGATGCTCACTCCCATAGATGGA 6166
Db 540 GATGATCCGGGCGCTGTGTCGGGAGCGGGGCGCCGCTCTGCTGACACGCGAGTACTCGA 599
QY 6167 GGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGTAATGGGGGTTTCCGTGCT 6226
Db 600 CGAGCGGACCACTTCGCGGACGAGCTGACGCTCATCGACACGCGCGCATCTGGCGCA 659
QY 6227 GGGCAGCCCGCAACATCTCAAGGCGAGATTGCGCGGGGTCAACATGACCTTCGGGT 6286
Db 660 GGGCACCCCGCGGAGCTGAAGGCGAGCGCGCGCGCGCGG---CGTGTCTGACGCTGCGGCT 716
QY 6287 GCCCGCGCAGGTCCAGCGCGGCGGCTTCTGTCGGCGCGGCTTCCCTGGTCCGA 6346
Db 717 GCGTGACCCCGAGCGCGGGCCGACGCGGGCGGCTTCTGTCGCAAGGCGCTGCGCGCGCGC 776
QY 6347 GCTGCGGAGGACATGGAGGCGCGCTTCCAGTGTGCGCGCGGAGGGCGCTGCGC 6406
Db 777 CGC---CGACTCGACTCCGATCCGGCGCGGCTGTGCTGGGTGACCGACCCGACCG 833
QY 6407 CTTGGCGCGGCTTTTGGAGAGCTGGCGGTGCAACGCGCGAGAGCACGCGCTGGAGGACTT 6466
Db 834 GGGCGGCTGCGCCCTGGGCGAGCTGGCGCGGGCGGCGCATCCA-----CGTCCGACGACTT 887
QY 6467 TTCCGTGAGCCAGCATGTGAGGAGGATTTCTTGTACTTCTTCAAGGACCGAGGAA 6526
Db 888 CACGCTGGCCAGCCCTCGCTCGACAGCGGTGTTCTTCGCGCTCACCGGTCTACTCGACGCT 947


```
Query Match      2.0%; Score 134.2; DB 3; Length 4403765;
Best Local Similarity 47.8%; Pred. No. 9.3e-18;
Matches 456; Conservative 0; Mismatches 488; Indels 9; Gaps 2;

Qy 5570 CACCAGGGGATGTTGGTCTGAGGAACTTGACCAAGGTATACCTGGGCGAGAGAT 5629
Db 3266492 CACACACACATGGCGGTGGTGTAAACGGGGTTCCGAAGACCTACGGCAAGGGCAAGAT 3266551

Qy 5630 GCCAGCTTTGACCGCTTGTGCTGGGGAATCCCTCTGCTGAGTGTGTTGGGCTGCTGGG 5689
Db 3266552 TGTGSCCTCGATGACGTGAGTTTCAAGTGGCGCGGTGAGTGTGAGTGTGCTGGG 3266611

Qy 5690 TGTGATGAGAGGAAAGACGTCAAGTTTGCATGCTGACGGGGACACATTTGGCCAG 5749
Db 3266612 CCCAAACGGGCGCGCAAGACCAATGCTGGACATCTTGTTCGACGCTGACCCGACCGGA 3266671

Qy 5750 CAGGGGAGCTGTGCTGGCAGGCCACAGCGTGGCCCCGGGACCCAGTGTGCGCACCT 5809
Db 3266672 TGCCGGCTCGCGCATCATCGTGGCTACGATGTTGTTTCCGAAACGGCGCGGTGACGCGG 3266731

Qy 5810 CAGCATGGGATATCGCCCTCAATCGATCGATCCATCTTTGAGCTGTGACGGGCGCGAGCA 5869
Db 3266732 CTCGATCATGTCCACGGGAGCAGGTGCGCTGACACACGCGCTTTCCGTGAGCAGAA 3266791

Qy 5870 CCTGAGCTGTGCTGCGCCCTGCGCGGTGTCCTCCGAGGCCACAGTTGCCAGACCGCTGG 5929
Db 3266792 CTTGCTGTGTTTGTGCTGCTGTGGGACTGAGCAAGTCCGCGCGCGCAACGCGCGC 3266851

Qy 5930 CTCAGGCTGCTGCTGCGGACTCTCATGTGTAACAGACCGGCTGAGGACACCTACAG 5989
Db 3266852 CGAATGCTGACGAAATTCAGCCCTGTAATGTCGCGGAAAGAGGGGGTGGACACCTATCTC 3266911

Qy 5990 CGGAGGAAACAAACGCAAGCTGGCGACCGCCCTGCGCTGTGTTGGGACCCAGCCCTGGT 6049
Db 3266912 CGGCGAATGCGCGACGAATAGACATCGCTGCGGATTTGTTGTTCCAAACCCAGGTGGC 3266971

Qy 6050 GTTTCTGGAACGAGCCGACACAGGCAATGGACCCAGCGCGCGGCTTCTCTTTGGAAACAG 6109
Db 3266972 GTTCTTAGACGAGCCCAACACGGGCTCGATCCCGAGGACCGCAAGCTATTTGGGATCT 3267031

Qy 6110 CTTTGGCCCTGTGCGGAGGGGCGCTTGTAGTATGCTCATCTCCATGACATGAGGAA 6169
Db 3267032 GGTGGCCAGCTTCAAGAAAGCTGGGCAATGGCCACGTTGTTGACCAACGCAATATCTCGAGGA 3267091

Qy 6170 GTGTGAACGCTCTGCTCGCCCTAGCCATCATGTTGATGAGTGGCGGCTTCCGCTGCTGG 6229
Db 3267092 GCGGATGCGCTCAGTGAACCGCATCATCTGATCGATCAACGCGATATCATCGCCGAGG 3267151

Qy 6230 CAGCCCGCAACATCTCAAGGGCAGATTGCGCGCGGCTCACAACATGACCCCTGCGGGTGGC 6289
Db 3267152 CACCGCAATGAACTCAAGCACCG-----CGCCGGCGACACCTTCTCGGAAATAGTGCC 3267205

Qy 6290 CGCCGAAAGTCCAGCCGAGCGGCTTCTGTTGGCGGCGGATTTCCCTGGGTGAGGACT 6349
Db 3267206 CCGCGATCTGAAGGATCTGAGCGCTATCGTCCGCGCGCTCGGTTCTGCTGTTGCGCGGAC 3267265

Qy 6350 GCGGAGGACATGAGGCGCCCTGCGCTTCCAGCTGCGCGGGAGGGCGCTGCGCCCT 6409
Db 3267266 CCAAGGCGATGCTGACGCCCGCATCAGACCGCATTCGATGCGCGCGCGCTGACGCGAT 3267325

Qy 6410 GGC---GCGGCTCTTTGGAGAGCTGGCGGTGCAACGCGCAGACGACGCGGTGGAGGACTT 6466
Db 3267326 AGTATGCTCTGAGGACGCGCGCGGATCGACGAGGAGGATCGAGTAGCCGATAT 3267385

Qy 6467 TTCCGTGAGCAGATGCTGGAGGAGGATTTCTTGTACTTCTTCCAAAGGACC 6519
Db 3267386 TGCCTGCGCGGACCGTCACTCGATCAGTATTCTTGGCCATGACGACCGGATC 3267438
```

RESULT 13

US-09-103-840A-1

; Sequence 1, Application US/09103840A

```
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1
```

```
Query Match      2.0%; Score 134.2; DB 3; Length 441529;
Best Local Similarity 47.8%; Pred. No. 9.3e-18;
Matches 456; Conservative 0; Mismatches 488; Indels 9; Gaps 2;

Qy 5570 CACCAGGGGATGTTGGTCTGAGGAACTTGACCAAGGTATACCTGGGCGAGAGAT 5629
Db 3272216 CACACACACATGGCGGTGGTGTAAACGGGGTTCCGAAGACCTACGCAAGGGCAAGAT 3272275

Qy 5630 GCCAGCTTTGACCGCTTGTGCTGGGGAATCCCTCTGCTGAGTGTGTTGGGCTGCTGGG 5689
Db 3272276 TGTGSCCTCGATGACGTGAGTTTCAAGGTGGCGCGGTGAGTGTGCTGGGCTGCTGGG 3272335

Qy 5690 TGTCAATGGAGCAGGGAAGACGTCACGTTTTCGATGCTGACGGGGACACATTTGGCCAG 5749
Db 3272336 CCCCAACGGGCGCGCAGACGACATGTTGCGACGCTGACCCGACCGGA 3272395

Qy 5750 CAGGGGCGAGCTGTGCTGCGAGCCACAGCGTGGCCCCGGGAAACCCAGTGTGCGCACCT 5809
Db 3272396 TGC CGCTGCGGATCATCGTGGCTACGATGTTGTTTCGAAACCGGCGGTGACGCG 3272455

Qy 5810 CAGCATGGGATATGCTCCCTCAATCCGATGCTGTTGAGTGTCTGACGCGCGCGGAGCA 5869
Db 3272456 CTCGATCATGTCAACCGGCGAGCAGGTGGCGCTGCGACGCGCTTTCGCTGAGCAGAA 3272515

Qy 5870 CTTGAGCTGCTTCCGCGCTGCGCGGTGTCGCGAGGCCAGGTGTCAGAGCCGCTGG 5929
Db 3272516 CTTGCTGTTGTTTGGTCTGCTGTTGGGACTGAGCAAGTCCGCGGCGCGAAACCGCGCGC 3272575

Qy 5930 CTCAGGCTGCGCGCTCTGGGACTCTCATGCTACGACAGCCGCTGCGAGGACCTACAG 5989
Db 3272576 CGAATGCTCGAGCAATTCAGCCCTCGTACATGCGGAAAGAGCGGGTGGGACCTACTC 3272635

Qy 5990 CGGAGGAAACAAACGCAAGCTTGGCGACCGCCCTGCGCGCTGTTGGGAGCCAGCCGCTGGT 6049
Db 3272636 CCGCGAATGCGCGCGACGAAATAGACATCGCTGCGGATTTGTTGTTCCAAACCCAGGTGGC 3272695

Qy 6050 GTTTCTGAGCAGCCGACACAGCAGCATGGACCCCGAGCGCGCGCTTCTCTTTGGAACAG 6109
Db 3272696 GTTCTTAGACGAGCCACACCGGCTCGATCCCGAGGCGCGCAAGCTATTTGGGATCT 3272755

Qy 6110 CTTTGGCGCTGTGCGGAGGGCGCTTTCAGTGTATGCTCACTCCATGACATGAGGAA 6169
Db 3272756 GTTGGCCAGCTTCAAGAACTGGGCAATTCACAGTGTGTTGACCAACGAGTATCTCGAGGA 3272815

Qy 6170 GTGTGAACGCTGTGCTGCGCCTTAGCCATCATGTTGAATGGGCGGTTCGCTGCTGCGG 6229
Db 3272816 GCGGATGCTGCTAGTGAACCGCATCATCTCTGATCGATCAACGCGATATCATCGCGAAGG 3272875

Qy 6230 CAGCCCGCAACATCTCAAGGGCAGATTTCGCGGCGGTCACACACTGACCCCTGCGGCTGCC 6289
Db 3272876 CACCGCAATGAACCTCAAGCACCG-----CGCCGGCGACACCTTCTTGGCAATAGTGCC 3272929

Qy 6290 CGCCGAAAGGTCCAGCGCGGACGCGCTTCTGTTGGCGCGGAGTTTCCCTGGTGGTGGAGCT 6349
```

Db 3272930 CCGGATCTGAAGATCTGACGCTATCTGTCGGCGCTCGGTTGCTGTTGCCGAGCA 3272989
Qy 6350 GCGGAGGACATGAGAGCGCGCTGCGCTTCAGCTGCGCGCGGAGGGCGCTGCGCCCT 6409
Db 3272990 CCACAGGGCGATGCTGACGCCCACTCAGACCGCATTTACGATCCGGCGCTGACGGCAT 3273049
Qy 6410 GGC---GCGGCTTTTGGAGAGTGGGGTGACGGGACAGACACGGCGTGGAGACTT 6466
Db 3273050 ACATATGCTCGTAGGACAGCGCGCGGATCGACGAGCGAGGATCGAGTTAGCCGATAT 3273109
Qy 6467 TTCGCTGAGCCAGACGATGCTGGAGGAGTATTCTTGTACTTCTTCAAAGGACC 6519
Db 3273110 TGGCTGCGCCGACCGCTCACTCGATCAGCTATTCTTGGCCATGACACCGATC 3273162

RESULT 14
US-09-252-991A-4588/c
; Sequence 4588, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4588
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4588

Query Match 1.8%; Score 125.2; DB 4; Length 999;
Best Local Similarity 48.5%; Pred. No. 5.5e-17;
Matches 377; Conservative 0; Mismatches 398; Indels 3; Gaps 1;

Qy 2676 GCCTCGGGGGCTCAGCTGGACTTCTACAGGGCCACATCACCGCTTCTCTGGGCGAC 2735
Db 939 GCCTCAAGGGCATCGACCTGGACGTCGCGAAGGTGATTTCTCGCTTGTCTGGCCCC 880
Qy 2736 AACGGGCGGGCAAGACCAACCCCTGTCCATCTTGAGTGGCCCTTTCACCCAGCTGGT 2795
Db 879 AACGGGCGGGCAAGTCCACCACTCGGATTTCTCGACCTTGGTGAACAGACACAGC 820
Qy 2796 GGTCTGCTTATCTTGGGCGAGGACGTCGCTCGATGCGGCGGCGGCGGCGGCGGCGG 2855
Db 819 GGTCTGCTTATCTTGGGCGAGGACGTCGCTCGATGCGGCGGCGGCGGCGGCGGCGG 760
Qy 2856 CTGGGCTGCTGCTCAGTACACGCTGCTTTGACATGCTGACGCTGGAGGAGACGTC 2915
Db 759 CTGGGCTGCTGCTGCGAGGAGTCACTTCAACAGTTTCAGAGGTTCTTGCATCGTC 700
Qy 2916 TGGTTTATGGGGGCTGAAAGGGTCTGAGTCCGCTGTAGTGGGCGGCTTCCAGGAGGACCGT 2975
Db 699 GTGACCCAGGCGGCTACTACGGCATCCCGGCGAAGATCCCAAGGAGCGCGCGAGCGC 640
Qy 2976 CTGCTCAGGATGTGGGGCTGGTCTCAACGACAGTGTGAGACTGCGCACTCTCTGGT 3035
Db 639 TACCTGACCCAACTGGGCGCTGTGGGACAGCGCAACGAAAGCTTCGCGGATGCTCTCGGC 580
Qy 3036 GGGATGCAACGGAAGCTGCTCGTGGCCATTTGCTTGTGGGCGGCTTCCAAAGTTGTTATC 3095
Db 579 GGGATGAGCGGCGCTGATGATGCTCCCGGCGACTGTTGTCACAGCGCGCTCTGATC 520
Qy 3096 CTGGAAGAGCTTACGGCTGGCGTGAATCTGCTTCCCGCGCGGATTTTGGGAGCTGCTG 3155
Db 519 CTCGACGAGCCCAACCGCGGGGTGACATCGAGCTGCTGTTTCGATGTGGAGCTTCTCTC 460

Qy 3156 CTCAAA---TACGGAAGGTCGACGCTGATCTCTCCACCACCACTGATGAGCA 3212
Db 459 ACAGAACTCAACACGGAAGGCATCAGATCATCTCTCACCACTTACTTGGAGAGGCG 400
Qy 3213 GAGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGCTGTGGCTCC 3272
Db 399 GAGCAGCTCTGCGGCAACATTGCCATCATCGACACGCGGAGATCGTGAGAACACAGC 340
Qy 3273 CCACTTTCTTGGCGCGTCACTGGGCTCGGCTACTACTGACGCTGTTGAGGGCGCGC 3332
Db 339 ATGCGCGACTCTGATGACCTTCTGCTCGGAAACCTTCTGCTCGACTGGAAGACGTC 280
Qy 3333 CTGCCCCCTGACCAACATGAGAGGCTGACATGACATGGAGGGCAGTGTGACACAGG 3392
Db 279 CAGGCCCTGCCGCGACCTCGACGGCTATCGACGGCTGTCAGACCACTCTC 220
Qy 3393 CAGGAAAGAAATGGCAGCGGACGAGTGGCACTCTCAGCTGTGGGCC 3450
Db 219 GAAGTCAGGTGGAGAGAGCCAGGCAACGACCTGTTCGCCCACTGGGTGCAC 162

RESULT 15
US-09-252-991A-4266
; Sequence 4266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4266
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4266

Query Match 1.8%; Score 125.2; DB 4; Length 1008;
Best Local Similarity 48.5%; Pred. No. 5.6e-17;
Matches 377; Conservative 0; Mismatches 398; Indels 3; Gaps 1;

Qy 2676 GCCTCGGGGGCTCAGCTGGACTTCTACAGGGCCACATCACCGCTTCTCTGGGCGAC 2735
Db 133 GCCTCAAGGGCATCGACCTGGACGTCGCGAAGGTGATTTCTCGCTTGTCTGGCCCC 192
Qy 2736 AACGGGCGGGCAAGACCAACCCCTGTCCATCTTGAGTGGCTTTCACCCAGCTGGT 2795
Db 193 AACGGGCGGGCAAGTCCACCACTCGGATTTCTCGACCTTGGTGAACAGACACAGC 252
Qy 2796 GGTCTGCTTATCTTGGGCGAGGAGTCTGAGTGGCGCTTCCAGATGCGGCGGCGGCGG 2855
Db 253 GGTCTGCTTATCTTGGGCGAGGAGTCTGAGTGGCGCTTCCAGATGCGGCGGCGGCGG 312
Qy 2856 CTGGGCGTCTGCTCAGTACCAACGCTGCTTTGACATGCTGACCTGGACGAGCAGCTC 2915
Db 313 CTGGGCGTGGTGGCGAGGAGTTCACCTTCAACAGGTTTCGAGAGGTTCTCGATCGTC 372
Qy 2916 TGGTTTATGGGCGGTGAAGGTCTGAGTGGCGCTGTTAGTGGGCGGCGGAGGAGCGGT 2975
Db 373 GTGACCCAGGCGGCTACTACGGCATCCCGGCGAAGATCCCAAGGAGCGCGCGAGCGC 432
Qy 2976 CTGCTCAGGATGTGGGGCTGGTCTCCAGCAGAGTGTGACAGTCTGCCACCTCTCTGTT 3035
Db 433 TACCTGACCCAACTGGGCGCTGTGGGACAGCGCAACGAAAGCTTCGCGGATGCTCTCGGC 492
Qy 3036 GGGATGCAACGGAAGCTGTCGCTGGGCGATTCCTTTGTGGGCGGCTCCCAAGTTGTTATC 3095

This Page Blank (Uspto)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	6804	100.0	6804	9	US-09-995-542-4	Sequence 4, Appli
2	6762.4	99.4	6791	16	US-10-332-447-59	Sequence 59, Appli
3	6717.8	98.7	6768	9	US-09-858-194-1	Sequence 1, Appli
4	6717.8	98.7	6768	15	US-10-154-419-1	Sequence 1, Appli
5	6710	98.6	7795	10	US-09-983-446A-8	Sequence 8, Appli
6	6686.4	98.3	6704	17	US-10-775-920-1	Sequence 1, Appli
7	6684.8	98.2	6704	17	US-10-775-920-3	Sequence 3, Appli
8	6565.6	96.5	6588	17	US-10-775-920-4	Sequence 4, Appli
9	6478.4	95.2	6522	16	US-10-182-006-1	Sequence 1, Appli
10	6398.4	94.0	6432	9	US-09-858-194-3	Sequence 3, Appli
11	6398.4	94.0	6432	15	US-10-154-419-3	Sequence 3, Appli
12	6073.6	89.3	6174	17	US-10-775-920-6	Sequence 6, Appli

ALIGNMENTS

```

US-09-995-542-4
; Sequence 4, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shuter, John
; APPLICANT: Uliasz, Larni
; TITLE OF INVENTION: ATP-Binding Ca
; FILE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,5
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (210)..(347)
; NAME/KEY: CDS
; LOCATION: (210)..(6650)
US-09-995-542-4

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Qy	61	AGCCCTGGGAAGAGCTTCCAGGAAACCTCTCGCTGTGGATAAAGGAATGAGGTTTCAGAAA	120
Db	61	AGCCCTGGGAAGAGCTTCCAGGAAACCTCTCGCTGTGGATAAAGGAATGAGGTTTCAGAAA	120
Qy	121	GGGACAGGAGTTGCCCGCAGCCGACGGACGTCCTTTCAGCCCGACGTTGTCTGTGACCT	180
Db	121	GGGACAGGAGTTGCCCGCAGCCGACGGACGTCCTTTCAGCCCGACGTTGTCTGTGACCT	180
Qy	181	CTCTGTCCCGTCCCGCCAGTCTCACCATGGGCTTCTGGACACAGCTGATGCTGCTGC	240
Db	181	CTCTGTCCCGTCCCGCCAGTCTCACCATGGGCTTCTGGACACAGCTGATGCTGCTGC	240
Qy	241	TCTGGAAAGAAATTTTCAATGATCGCCGGAGACAGCCGGTCCAGCTCTCTGGTCCAAATGCTGT	300
Db	241	TCTGGAAAGAAATTTTCAATGATCGCCGGAGACAGCCGGTCCAGCTCTCTGGTCCAAATGCTGT	300
Qy	301	GGCCTCTTCTCTTCTTCTATCTGTGTGGTGTTCGCCATCTCCACCCGCGCCTTGAGC	360
Db	301	GGCCTCTTCTCTTCTTCTATCTGTGTGGTGTTCGCCATCTCCACCCGCGCCTTGAGC	360
Qy	361	ACCATGATGCACTTCCCAAAAGAGCACTGCCATCGGGGGCACGTCGCTCGCTGGCTCC	420
Db	361	ACCATGATGCACTTCCCAAAAGAGCACTGCCATCGGGGGCACGTCGCTCGCTGGCTCC	420
Qy	421	AGGTCCTCATCTGTAATGTGAACAAACACTGCTTTTCGCAGCTGACACCGGGCGAGGAGC	480
Db	421	AGGTCCTCATCTGTAATGTGAACAAACACTGCTTTTCGCAGCTGACACCGGGCGAGGAGC	480
Qy	481	CCGGGCGCTGAGCAACTTTCAAACACTCTCTGTCTCCGGCTGCTAGCCGATGCCGCA	540
Db	481	CCGGGCGCTGAGCAACTTTCAAACACTCTCTGTCTCCGGCTGCTAGCCGATGCCGCA	540
Qy	541	CTGTGCTGGAGGGGCCAGTCCCCACAGGACGCTGGCTGGGCTTAGGAAAGCTGATCGCCA	600
Db	541	CTGTGCTGGAGGGGCCAGTCCCCACAGGACGCTGGCTGGGCTTAGGAAAGCTGATCGCCA	600
Qy	601	CGCTGAGGGCTGACGACGACGCGCCAGCCTCAAACCAACCAAGCAGCTCTCCATGGAAC	660
Db	601	CGCTGAGGGCTGACGACGACGCGCCAGCCTCAAACCAACCAAGCAGCTCTCCATGGAAC	660
Qy	661	CACCATGTGGATGTTCGGAGTGTGTGAGCTCATCTGTGCGACCGGAATTCCTTGGGGT	720
Db	661	CACCATGTGGATGTTCGGAGTGTGTGAGCTCATCTGTGCGACCGGAATTCCTTGGGGT	720
Qy	721	TGGCACTGGGCCAAGCCAGAGCCCTTTCACACGCTTGTGGAGGCGCTGAGGACCTGG	780
Db	721	TGGCACTGGGCCAAGCCAGAGCCCTTTCACACGCTTGTGGAGGCGCTGAGGACCTGG	780
Qy	781	CCCAGGAGCTCTCGCGCTGCGCAGCCTGTGGAGCTTCGGGCACTGCTGCAGAGACCCC	840
Db	781	CCCAGGAGCTCTCGCGCTGCGCAGCCTGTGGAGCTTCGGGCACTGCTGCAGAGACCCC	840
Qy	841	GAGGGACAGGGCCCTGTGAGTTGTCTGAGAGCCCTCTGCAGTGTTCAGGGGACCTA	900
Db	841	GAGGGACAGGGCCCTGTGAGTTGTCTGAGAGCCCTCTGCAGTGTTCAGGGGACCTA	900
Qy	901	GCAGCAGATGGGGCCCTCCCTCACTGTTACAGAGGCTAGTGACCTCAAGAGCTGTGG	960
Db	901	GCAGCAGATGGGGCCCTCCCTCACTGTTACAGAGGCTAGTGACCTCAAGAGCTGTGG	960
Qy	961	GGCAGGAGCCAGAAATCCGCCCTGCGACAGCAGCCTGAGCCCCGCTGTCTCGGAGCTGA	1020
Db	961	GGCAGGAGCCAGAAATCCGCCCTGCGACAGCAGCCTGAGCCCCGCTGTCTCGGAGCTGA	1020
Qy	1021	TTGAGGCCCTGGAACAGCCACCCGCTGTCCCGCTGTCTTGGAGACGCTGAAAGCTCTGA	1080
Db	1021	TTGAGGCCCTGGAACAGCCACCCGCTGTCCCGCTGTCTTGGAGACGCTGAAAGCTCTGA	1080
Qy	1081	TCCTCGGGAGCTACTCTTTTGACACAGATACACCTTTTACCCGAAAGCTCATGSCCCAGG	1140
Db	1081	TCCTCGGGAGCTACTCTTTTGACACAGATACACCTTTTACCCGAAAGCTCATGSCCCAGG	1140
Qy	1141	TGAAACCGGACCTTCAGGAGCTCACCCCTGTGAGGGATGTCCGGGAGGTTGGGAGATGC	1200

[illegible]

QY	4441	AGGTACAGTACGAGGCTTCTCGTGGGGGCGAGACCCAGGCTGCGCTCGGGCAAG	4500
Db	4441	AGGTACAGTACGAGGCTTCTCGTGGGGGCGAGACCCAGGCTGCGCTCGGGCAAG	4500
QY	4501	AGTTGGGCGCTCAGTGGAGAGTTGTGGCGCTGCTGAGTCCCTGCTGGGGGGCCC	4560
Db	4501	AGTTGGGCGCTCAGTGGAGAGTTGTGGCGCTGCTGAGTCCCTGCTGGGGGGCCC	4560
QY	4561	TCGACCGTGTCTGAAACCTCAGAGCTGGGGTCAACGCTGGATGCTCAGGACATC	4620
Db	4561	TCGACCGTGTCTGAAACCTCAGAGCTGGGGTCAACGCTGGATGCTCAGGACATC	4620
QY	4621	TCAGATCTGGTTCAACAAAGAGCTGGCACTCCATGGTGGCTTGTCAACCGAGCCA	4680
Db	4621	TCAGATCTGGTTCAACAAAGAGCTGGCACTCCATGGTGGCTTGTCAACCGAGCCA	4680
QY	4681	GCAACGCAATCTCTCGGTGCTCACTGCCCCAGGCGCGGCCCGCCACAGCATCA	4740
Db	4681	GCAACGCAATCTCTCGGTGCTCACTGCCCCAGGCGCGGCCCGCCACAGCATCA	4740
QY	4741	CCACATCAACCCCTTGAAACCTCAACGAGGAGCTGTCTGAGGCTGCACTGATGG	4800
Db	4741	CCACATCAACCCCTTGAAACCTCAACGAGGAGCTGTCTGAGGCTGCACTGATGG	4800
QY	4801	CTCTCGGTGGAGCTCTCGTCTCCATCTGTGTGCTTTGGCATGCTCTTGTCCCGG	4860
Db	4801	CTCTCGGTGGAGCTCTCGTCTCCATCTGTGTGCTTTGGCATGCTCTTGTCCCGG	4860
QY	4861	CAAGCTTCACTCTTGTCTCACTGAGGAGGAGTCAACCCAGGCAAGCACTCGAGCTCA	4920
Db	4861	CAAGCTTCACTCTTGTCTCACTGAGGAGGAGTCAACCCAGGCAAGCACTCGAGCTCA	4920
QY	4921	TGGGGGCTGTGCCCCACCTCTA CTGGCTTGGGAACTTTCTCTGGGA CATGTAACT	4980
Db	4921	TGGGGGCTGTGCCCCACCTCTA CTGGCTTGGGAACTTTCTCTGGGA CATGTAACT	4980
QY	4981	ACTTGTGCCAGATGATCGTGGTCTCATCTTCTGGGCTTCCAGCAGAGGCAATG	5040
Db	4981	ACTTGTGCCAGATGATCGTGGTCTCATCTTCTGGGCTTCCAGCAGAGGCAATG	5040
QY	5041	TGGCCCTGCAACCTGCTCTCTCTGCTGTGTGCTACTGATGCTGCTCGATCA	5100
Db	5041	TGGCCCTGCAACCTGCTCTCTCTGCTGTGTGCTACTGATGCTGCTCGATCA	5100
QY	5101	CACGCTCATGATCCAGGCTCTCTCTCTCGTGGCCAGCAAGCCCTATGTGGTGC	5160
Db	5101	CACGCTCATGATCCAGGCTCTCTCTCTCGTGGCCAGCAAGCCCTATGTGGTGC	5160
QY	5161	TCACTGCATAAACCTCTTTATGGCATCAATGGAAGCATGGCCACCTTTGTGCTTGAGC	5220
Db	5161	TCACTGCATAAACCTCTTTATGGCATCAATGGAAGCATGGCCACCTTTGTGCTTGAGC	5220
QY	5221	TCCTTCTGATCAGAACTCAGGAGGTGAGCCGGATCTTTGAAACAGGTCTTCTTATCT	5280
Db	5221	TCCTTCTGATCAGAACTCAGGAGGTGAGCCGGATCTTTGAAACAGGTCTTCTTATCT	5280
QY	5281	TCGCCCACTTCTGCTTGGGCGGGGCTCATATGATGCTGGAAACCGAGCCATGGCTG	5340
Db	5281	TCGCCCACTTCTGCTTGGGCGGGGCTCATATGATGCTGGAAACCGAGCCATGGCTG	5340
QY	5341	ATGCCCTTGGAGCGCTTGGGAGACAGGAGTCCAGTCAACCCCTGCGCTGGAGGTGCTG	5400
Db	5341	ATGCCCTTGGAGCGCTTGGGAGACAGGAGTTCAGTCAACCCCTGCGCTGGAGGTGCTG	5400
QY	5401	GCAAGAACTCTTGGCCATGGTGAATAAGGGGCGCTCTTCTTCTTCACTACTG	5460
Db	5401	GCAAGAACTCTTGGCCATGGTGAATAAGGGGCGCTCTTCTTCTTCACTACTG	5460
QY	5461	TGAGCAACCAAGCAACTCTCTGCAACAGCCAGGCTGAGGTCTCTGCACTCTCGGGAG	5520
Db	5461	TGAGCAACCAAGCAACTCTCTGCAACAGCCAGGCTGAGGTCTCTGCACTCTCGGGAG	5520
QY	5521	AGGAGNACGAGGATGTAGCCGTGACGGGAGCGGGTGGTCCAGGAGCCACCCAGGGG	5580

Db	5521	AGGAGNACGAGGATGTAGCCGTGAAACGGGAGCGGTGGTCCAGAGGCCACCCAGGGG	5580
QY	5581	ATGTGTTGTGCTGAGGAACTTGACCAAGGTATACCGTGGGAGAGGATGCGAGCTGTTG	5640
Db	5581	ATGTGTTGTGCTGAGGAACTTGACCAAGGTATACCGTGGGAGAGGATGCGAGCTGTTG	5640
QY	5641	ACCGCTTGTGCTGGGATTTCCCTGTGTAGTGTGTTTGGGCTGCTGGGTGTAATGGAG	5700
Db	5641	ACCGCTTGTGCTGGGATTTCCCTGTGTGTAGTGTGTTTGGGCTGCTGGGTGTAATGGAG	5700
QY	5701	CAGGGAAGACGTCCACGTTTGCATGCTGACCGGGGACACATTGGGCAGCAGGGGCGAGG	5760
Db	5701	CAGGGAAGACGTCCACGTTTGCATGCTGACCGGGGACACATTGGGCAGCAGGGGCGAGG	5760
QY	5761	CTGTGCTGGAGGCAACAGCGTGGCCCGGAAACCCAGTGTCTGCGACCTCAGATGGGAT	5820
Db	5761	CTGTGCTGGAGGCAACAGCGTGGCCCGGAAACCCAGTGTCTGCGACCTCAGATGGGAT	5820
QY	5821	ACTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGCGAGCACCTGGAGCTGC	5880
Db	5821	ACTGCCCTCAATCCGATGCCATCTTTGAGCTGTGACGGGCCGCGAGCACCTGGAGCTGC	5880
QY	5881	TTGCGCGCTCGCGGTGTCCCGAGGCGCCAGGTTGCCAGACCGCTGGCTCAGGGCTGG	5940
Db	5881	TTGCGCGCTCGCGGTGTCCCGAGGCGCCAGGTTGCCAGACCGCTGGCTCAGGGCTGG	5940
QY	5941	CGCGTCTGGAGCTCTCATGTACGAGACCGGCTCGAGGCACTTACAGCGAGGGAACA	6000
Db	5941	CGCGTCTGGAGCTCTCATGTACGAGACCGGCTCGAGGCACTTACAGCGAGGGAACA	6000
QY	6001	AACGCAAGCTGGCGACCGGCTGGTGGGACCGGCTGGTGGTCTGAGCTG	6060
Db	6001	AACGCAAGCTGGCGACCGGCTGGTGGGACCGGCTGGTGGTCTGAGCTG	6060
QY	6061	AGCCGACCAAGGATGAGCCCGCGCTTCTTTGGAAACAGCTTTTGGCCG	6120
Db	6061	AGCCGACCAAGGATGAGCCCGCGCTTCTTTGGAAACAGCTTTTGGCCG	6120
QY	6121	TGCTGCGGAGGCGCTTCACTGATGCTCACTCCATAGATGAGAGGATGTGAAGCGC	6180
Db	6121	TGCTGCGGAGGCGCTTCACTGATGCTCACTCCATAGATGAGAGGATGTGAAGCGC	6180
QY	6181	TCCTGCTCGGCTTACGATCATGAGTGGGCGGTTCGCTGCTGGGAGCGCCGCAAC	6240
Db	6181	TCCTGCTCGGCTTACGATCATGAGTGGGCGGTTCGCTGCTGGGAGCGCCGCAAC	6240
QY	6241	ATCTCAAGGCGAGATTCGGGCGGGTCAACACTGACCTTGCCTGCGGGTGCCTGCGAGGT	6300
Db	6241	ATCTCAAGGCGAGATTCGGGCGGGTCAACACTGACCTTGCCTGCGGGTGCCTGCGAGGT	6300
QY	6301	CCAGCCGCGAGCGCTTCTGCTGGCGCGAGTTCCTGCTGGTGGAGCTGCGGAGGAC	6360
Db	6301	CCAGCCGCGAGCGCTTCTGCTGGCGCGAGTTCCTGCTGGTGGAGCTGCGGAGGAC	6360
QY	6361	ATGGAAGCGCTGCGCTTCCAGCTGCGCGCGGAGGCGCTGCGCCCTGGCGCGCTCT	6420
Db	6361	ATGGAAGCGCTGCGCTTCCAGCTGCGCGCGGAGGCGCTGCGCCCTGGCGCGCTCT	6420
QY	6421	TTGGAGAGCTGGCGGTGCAACCGCGTGGAGGACTTTTCCGTGAGCGCAGA	6480
Db	6421	TTGGAGAGCTGGCGGTGCAACCGCGTGGAGGACTTTTCCGTGAGCGCAGA	6480
QY	6481	CGATGCTGGAGGAGTATCTTGTATCTTCTCAAGGACCAAGGGGAGGAGCAGACCG	6540
Db	6481	CGATGCTGGAGGAGTATCTTGTATCTTCTCAAGGACCAAGGGGAGGAGCAGACCG	6540
QY	6541	AAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCGCAGGCTGCGACACCCAAAC	6600
Db	6541	AAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCGCAGGCTGCGACACCCAAAC	6600
QY	6601	GCCTCAGCAGTTCCTCGATGACCTAGCACTGCCGAGACTGTGCTCTGAGCTTCCCTCC	6660

Db	1260	CAGGATGAAGAAAGGACAGCCAGACTTGGAGCGCGACACATGGAGGCGCTTGGCA	1319
Qy	1332	TCCTTTCTGACACCTTGGAGCGGTGGCTACAGCTGGCAGGACGGACACACGCTGATGTGGG	1391
Db	1320	TCCTTTCTGACACCTTGGAGCGGTGGCTACAGCTGGCAGGACGGACACACGCTGATGTGGG	1379
Qy	1392	CACCTGGTGGGACGCTGGGCGCGAGTGA CGGAGTGCCTGTCTTGGACAAGCTGGAGGGC	1451
Db	1380	CACCTGGTGGGACGCTGGGCGCGAGTGA CGGAGTGCCTGTCTTGGACAAGCTGGAGGGC	1439
Qy	1452	GCACCTCTCAGAGGACGCTTGGTGTCTCGCGGCGCTGCAACTGTCTCGCGGAACATTCGATTC	1511
Db	1440	GCACCTCTCAGAGGACGCTTGGTGTCTCGCGGCGCTGCAACTGTCTCGCGGAACATTCGATTC	1499
Qy	1512	TGGGCGGCGTGTCTTCTTGGGACCTGAGGA CTCTTCAACCCCAAGACACCCCAACC	1571
Db	1500	TGGGCGGCGTGTCTTCTTGGGACCTGAGGA CTCTTCAACCCCAAGACACCCCAACC	1559
Qy	1572	CCAGACCTGGGCGCGGACGCTGGCATCAAAATCCGCATGACATTTGACGTGGTCAAG	1631
Db	1560	CCAGACCTGGGCGCGGACGCTGGCATCAAAATCCGCATGACATTTGACGTGGTCAAG	1619
Qy	1632	AGGACCAATTAAGATCAGGACAGGTTTGGGACCTTGGGCGGACCCCTGACCC	1691
Db	1620	AGGACCAATTAAGATCAGGACAGGTTTGGGACCTTGGGCGGACCCCTGACCC	1679
Qy	1692	GACCTGGCGTACGTGTGGGCGGCTTCGTGTACCTGCAAGACCTGTGTGGAGCGTGCAGCC	1751
Db	1680	GACCTGGCGTACGTGTGGGCGGCTTCGTGTACCTGCAAGACCTGTGTGGAGCGTGCAGCC	1739
Qy	1752	GTGGGCTGTCTAGCGGCGCAACCCCGGCGGCGCTCTACCTGCAGACAGATGCCTAT	1811
Db	1740	GTGGGCTGTCTAGCGGCGCAACCCCGGCGGCGCTCTACCTGCAGACAGATGCCTAT	1799
Qy	1812	CCGTGCTATGTGACGACGCTTCTCGTGTGTCTGAGCGGCTGCTGCCGCTCTTCCCTG	1871
Db	1800	CCGTGCTATGTGACGACGCTTCTCGTGTGTCTGAGCGGCTGCTGCCGCTCTTCCCTG	1859
Qy	1872	ACGTGGCGCTGGATCTACTCCGTGACACTGACAGTGAAGCGCGTGTGGCGGGAAGAG	1931
Db	1860	ACGTGGCGCTGGATCTACTCCGTGACACTGACAGTGAAGCGCGTGTGGCGGGAAGAG	1919
Qy	1932	ACGGGCTGGGACACCATGCGGCGATGGGCTCAGCGCGGCTCTCGGCTAGGC	1991
Db	1920	ACGGGCTGGGACACCATGCGGCGATGGGCTCAGCGCGGCTCTCGGCTAGGC	1979
Qy	1992	TGGTTCTCAGCTGCGCTCGGCGCGCTTCTGCTCAGCGCGGCTGCTGTCTGTGGTCTC	2051
Db	1980	TGGTTCTCAGCTGCGCTCGGCGCGCTTCTGCTCAGCGCGGCTGCTGTCTGTGGTCTC	2039
Qy	2052	AAGCTGGGGACATCTCTCCCTACAGCCA CCGGGCGTGTCTTCTGTGTTGGCAGCC	2111
Db	2040	AAGCTGGGGACATCTCTCCCTACAGCCA CCGGGCGTGTCTTCTGTGTTGGCAGCC	2099
Qy	2112	TTGGCGGTGGCCAGCGTGAACCCAGAGCTTCGTGCTCAGCGCGCTTCTTCTCCGCGCAAC	2171
Db	2100	TTGGCGGTGGCCAGCGTGAACCCAGAGCTTCGTGCTCAGCGCGCTTCTTCTCCGCGCAAC	2159
Qy	2172	CTGGCTCGGCGCTTGGCGGCGCTTGGCTCTGCTGCTACCTGCTGCTGCTGCTGCTGCT	2231
Db	2160	CTGGCTCGGCGCTTGGCGGCGCTTGGCTCTGCTGCTACCTGCTGCTGCTGCTGCTGCT	2219
Qy	2232	GTGGCTTGGCGGACCGGCTGCGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGT	2291
Db	2220	GTGGCTTGGCGGACCGGCTGCGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGT	2279
Qy	2292	GTGGCTTGGCGGTGGCGGCGGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2351
Db	2280	GTGGCTTGGCGGTGGCGGCGGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2339
Qy	2352	CAGTGGCAACAGTGGGCAACCGGCGCTACGCGACAGCTCTTACGCTTGGCGCGGCTCT	2411

Db	2340	CAGTGGCAACAGTGGGCAACCGGCGCTACGCGACAGCTCTTACGCTTGGCGCGGCTCTCT	2399
Qy	2412	GGCTTCTGCTGCTGAGCGGGCGCTCTACGGGCTCGCACCTGGTACCTGGAGCTGTG	2471
Db	2400	GGCTTCTGCTGCTGAGCGGGCGCTCTACGGGCTCGCACCTGGTACCTGGAGCTGTG	2459
Qy	2472	TGCCAGGCGCAGTACGGGATCCCTGAACCATGGAATTTTCTTTTTCGGAGGAGCTACTGG	2531
Db	2460	TGCCAGGCGCAGTACGGGATCCCTGAACCATGGAATTTTCTTTTTCGGAGGAGCTACTGG	2519
Qy	2532	TGCGGACCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCCAACCCCGTGGACCCAAAGGTG	2591
Db	2520	TGCGGACCTCGGCCCCCAAGAGTCCAGCCCTTGGCCCCCAACCCCGTGGACCCAAAGGTG	2579
Qy	2592	CTGGTAGAAGAGGACACCGCGCGGCTGAGTCTGGGCTATCCGTTTCCGAGCCTGGAGAAG	2651
Db	2580	CTGGTAGAAGAGGACACCGCGCGGCTGAGTCTGGGCTATCCGTTTCCGAGCCTGGAGAAG	2639
Qy	2652	CGCTTCTTGGGAAGCCCGCAGCCCTGCGGGGCTCAGCGCTGAGCTTCTACCAAGGCG	2711
Db	2640	CGCTTCTTGGGAAGCCCGCAGCCCTGCGGGGCTCAGCGCTGAGCTTCTACCAAGGCG	2699
Qy	2712	CACATCACCGCTTCTTGGGCGCACAA CGGGGCGCGCAAGACCAACCTGTCCATCTTG	2771
Db	2700	CACATCACCGCTTCTTGGGCGCACAA CGGGGCGCGCAAGACCAACCTGTCCATCTTG	2759
Qy	2772	AGTGGGCTCTTCCACCCAGTGGTGTCTGCTGCTTCACTCTGGGCGCAGACGTCGCTCC	2831
Db	2760	AGTGGGCTCTTCCACCCAGTGGTGTCTGCTGCTTCACTCTGGGCGCAGACGTCGCTCC	2819
Qy	2832	AGCATGGCGCGCATCCGGGCGCCACCTGGGCGCTGTCTCTCAGTACAAAGTGTGTTGAC	2891
Db	2820	AGCATGGCGCGCATCCGGGCGCCACCTGGGCGCTGTCTCTCAGTACAAAGTGTGTTGAC	2879
Qy	2892	ATGCTGACCGTGGACGAGCA CGTCTGGTTCCTATGGGCGGCTGAAGGTCTGAGTGCCT	2951
Db	2880	ATGCTGACCGTGGACGAGCA CGTCTGGTTCCTATGGGCGGCTGAAGGTCTGAGTGCCT	2939
Qy	2952	GTAGTGGGCGCGGACGAGCACCGTCTGTCTGAGGATGTGGGCTGTCTTCCAGACAGT	3011
Db	2940	GTAGTGGGCGCGGACGAGCACCGTCTGTCTGAGGATGTGGGCTGTCTTCCAGACAGT	2999
Qy	3012	GTGACAGCTCGCCACCTCTCTGTGGGATGCAACCGGAAGCTGCTCCGTGGCCATTTGCTTT	3071
Db	3000	GTGACAGCTCGCCACCTCTCTGTGGGATGCAACCGGAAGCTGCTCCGTGGCCATTTGCTTT	3059
Qy	3072	GTGGGCGGCTCCCAAGTTGTTATCTTGA CGAGCCTACCGGCTGGCGTGTGATCTGTCTCC	3131
Db	3060	GTGGGCGGCTCCCAAGTTGTTATCTTGA CGAGCCTACCGGCTGGCGTGTGATCTGTCTCC	3119
Qy	3132	CGCGCGGTATTTGGAGCTGTGCTCAAAATACCGAAGGTCGACGCTGATCTCTCTCC	3191
Db	3120	CGCGCGGTATTTGGAGCTGTGCTCAAAATACCGAAGGTCGACGCTGATCTCTCTCC	3179
Qy	3192	ACCCACCACTGATGACGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGC	3251
Db	3180	ACCCACCACTGATGACGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGC	3239
Qy	3252	CGCTTGTGCTGTGGCTCCCCACCTTCTTCTGCGCGCTCACCTGGGCTCCCGCTACTAC	3311
Db	3240	CGCTTGTGCTGTGGATCCCCACCTTCTTCTGCGCGCTCACCTGGGCTCCCGCTACTAC	3299
Qy	3312	CTGACGCTGTGAGGCGCGCTTGGCGCTTGA CCAACCAATGAGAGGCTGACATGACATG	3371
Db	3300	CTGACGCTGTGAGGCGCGCTTGGCGCTTGA CCAACCAATGAGAGGCTGACATGACATG	3359
Qy	3372	GAGGCGAGTGTGGACACCGGCAAGGAAAGAAATGGCAGCGCAGGCGCAGCAGTCCGC	3431
Db	3360	GAGGCGAGTGTGGACACCGGCAAGGAAAGAAATGGCAGCGCAGGCGCAGCAGTCCGC	3419
Qy	3432	ACTCTCAGCTGTGGCGCTTGGTACAGCA CTGCGGTGCCCGGCGCA CGGCTGTGGAGGAG	3491
Db	3420	ACTCTCAGCTGTGGCGCTTGGTACAGCA CTGCGGTGCCCGGCGCA CGGCTGTGGAGGAG	3479

[illegible]

Db	6720	CTGCGAGGGGTGGTGCCCTGGAGAAAATAAAGAAGGCTGGAGAGAAAGCCGTGGTGTG
Qy	6792	GTGAATAAAAAA 6803
Db	6780	GTGAATAAAAAA 6791
RESULT 3		
US-09-858-194-1		
; Sequence 1, Application US/09858194		
; Patent No. US2002061590A1		
; GENERAL INFORMATION:		
; APPLICANT: GLUCKSMANN, MARIA		
; APPLICANT: CURTIS, RORY A.J.		
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF		
; FILE REFERENCE: WNI-153		
; CURRENT APPLICATION NUMBER: US/09/858,194		
; CURRENT FILING DATE: 2001-05-14		
; PRIOR APPLICATION NUMBER: 60/204,211		
; PRIOR FILING DATE: 2000-05-12		
; NUMBER OF SEQ ID NOS: 3		
; SOFTWARE: PatentIn Ver. 2.0		
; SEQ ID NO 1		
; LENGTH: 6768		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
; FEATURE:		
; NAME/KEY: CDS		
; LOCATION: (165)..(6596)		
US-09-858-194-1		
Query Match 98.7%; Score 6717.8; DB 9; Length 6768;		
Best Local Similarity 99.7%; Pred. No. 0;		
Matches 6741; Conservative 0; Mismatches 12; Indels 6; Gaps 1;		
Qy	46	CGCAAGAGCTCGAGACCCCCTGGAAGAGTTCCAGGAACCCCTGCCTGTGGGATTAAGG 105
Db	1	CGCAAGAGCTCGAGACCCCCTGGAAGAGTTCCAGGAACCCCTGCCTGTGGGATTAAGG 60
Qy	106	AATGAGTTTCAGAAAGGGCAGGAGTTGCCCGACGCCGACCGCACGCTTTCAGCCCCGA 165
Db	61	AATGAGTTTCAGAAAGGGCAGGAGTTGCCCGACGCCGACCGCACGCTTTCAGCCCCGA 120
Qy	166	CCGTTGCTTCGACTCTCTCTGCCCGTCCCGGCCAGTCTCACATGGCCCTTCTGGACAC 225
Db	121	CCGTTGCTTCGACTCTCTCTGCCCGTCCCGGCCAGTCTCACATGGCCCTTCTGGACAC 180
Qy	226	AGCTGATGCTGCTCTGGAAGAAATTTCATATATCGCCGAGACACCAGGTCACGCTCC 285
Db	181	AGCTGATGCTGCTCTGGAAGAAATTTCATATATCGCCGAGACACCAGGTCACGCTCC 234
Qy	286	TGCTGCAATTTGCTGTGGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 345
Db	235	TGCTGCAATTTGCTGTGGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 294
Qy	346	ACCGCCCTCGAGACCATGAATGACATTTCCCAAACAAGACCACTGCCATCGCGGGCA 405
Db	295	ACCGCCCTCGAGACCATGAATGACATTTCCCAAACAAGACCACTGCCATCGCGGGCA 354
Qy	406	CCGTGCCCTCGCTCCAGGGTCTCATCTGTAATGTGAACAAACACTGCTTTCCGACGCTGA 465
Db	355	CCGTGCCCTCGCTCCAGGGTCTCATCTGTAATGTGAACAAACACTGCTTTCCGACGCTGA 414
Qy	466	CACGGGCGAGGACCGCGCGCTGAGCAACTTCAACGACTCCCTGGTCTCCGGCTGC 525
Db	415	CACGGGCGAGGACCGCGCGCTGAGCAACTTCAACGACTCCCTGGTCTCCGGCTGC 474
Qy	526	TAGCCGATGCCGACCTGTCTGGGAGGGGCCAGTGCACAGAGACGCTGGCTGGCTAG 585
Db	475	TAGCCGATGCCGACCTGTCTGGGAGGGGCCAGTGCACAGAGACGCTGGCTGGCTAG 534
Qy	586	GGAAGCTGATCGGCACGCTGAGGGCTGCACGCGACGCGGCCACGCTCAACCAACCAAGC 645

2806 TCATCTGGGCCACGAGCTCGCTCCAGATGCGCGCCATCCGCGCCCACTCGGGGCTCT 2865
2755 TCATCTGGGCCACGAGCTCGCTCCAGATGCGCGCCATCCGCGCCCACTCGGGGCTCT 2814
2866 GTCTCTAGTACAAAGTCTGTTGACATGTCGACCGTGGACGAGACGTCGTCTTATG 2925
2815 GTCTCTAGTACAAAGTCTGTTGACATGTCGACCGTGGACGAGACGTCGTCTTATG 2874
2926 GCGGCTGAGGCTGCTGAGTGGCGCTGTAGTGGGCCCCGAGCAGGACCGTCTGTCAGG 2985
2875 GCGGCTGAGGCTGCTGAGTGGCGCTGTAGTGGGCCCCGAGCAGGACCGTCTACTGAGG 2934
2986 ATGTGGGCTTGGTCTCCAAAGCAGAGTGTGACAGTCCGCCACCTCTCTGTGGGATGCAAC 3045
2935 ATGTGGGCTTGGTCTCCAAAGCAGAGTGTGACAGTCCGCCACCTCTCTGTGGGATGCAAC 2994
3046 GGAAGCTGTCTGCTGGCCATTTGCTTGTGGGCGCTCCAAAGTGTGTTATCCTGTGAGCAG 3105
2995 GGAAGCTGTCTGCTGGCCATTTGCTTGTGGGCGCTCCAAAGTGTGTTATCCTGTGAGCAGC 3054
3106 CTAGCGCTGGGCTGGATCTCTCCCGCGCGGTATTTGGGAGCTGCTCTCAATATCC 3165
3055 CTAGCGCTGGGCTGGATCTCTCCCGCGCGGTATTTGGGAGCTGCTCTCAATATCC 3114
3166 GAGAAGTCTGACGCTGATCTCTCCACCCACACCTGGATGAGGACAGCTGCTGGAG 3225
3115 GAGAGTCTGACGCTGATCTCTCCACCCACACCTGGATGAGGACAGCTGCTGGAG 3174
3226 ACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGTGCTGCTGCGCTCCCACTCTTCTGTC 3285
3175 ACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGTGCTGCTGCGCTCCCACTCTTCTGTC 3234
3286 GCGGTACCTGGGCTCGGCTACTACTGAGCTGTGTGAAGCGCGCTGCGCTCCCTGACCA 3345
3235 GCGGTACCTGGGCTCGGCTACTACTGAGCTGTGTGAAGCGCGCTGCGCTCCCTGACCA 3294
3346 CCAATGAGAAGCTGACACTGACATGAGGAGGAGTGTGGACACCCAGGACGAGAAAGAGA 3405
3295 CCAATGAGAAGCTGACACTGACATGAGGAGGAGTGTGGACACCCAGGACGAGAAAGAGA 3354
3406 ATGGACCCAGGCGACGAGTGGGCACTCTCAGCTGCTGTGGCTGGCTGGTACAGCACTGGG 3465
3355 ATGGACCCAGGCGACGAGTGGGCACTCTCAGCTGCTGTGGCTGGTACAGCACTGGG 3414
3466 TGCCCGGGGACGCGCTGGTGGAGAGTGTGCACACGAGCTGTGTGCTGCTGCGCTTACA 3525
3415 TGCCCGGGGACGCGCTGGTGGAGAGTGTGCACACGAGCTGTGTGCTGCTGCGCTTACA 3474
3526 CCGGTGCCATGACGCGAGCTTCGCCACACTCTTTCGAGAGCTAGACACGCGCTGGCGG 3585
3475 CCGGTGCCATGACGCGAGCTTCGCCACACTCTTTCGAGAGCTAGACACGCGCTGGCGG 3534
3586 AGCTGAGGCTCACTGGCTACGGATCTCGACACCGACCTCGAGGAGATCTTCTGAAGG 3645
3535 AGCTGAGGCTCACTGGCTACGGATCTCGACACCGACCTCGAGGAGATCTTCTGAAGG 3594
3646 TGGTGGAGGAGTGTGCTGCGGACACAGATATGAGGAGTGGCAGCTGGCGGACGACCTAT 3705
3595 TGGTGGAGGAGTGTGCTGCGGACACAGATATGAGGAGTGGCAGCTGGCGGACGACCTAT 3654
3706 GCACAGGATTTGCTGGCTTAGACGTAACTCTGCGGCTCAAGATGTCGCGCCACAGAGACAG 3765
3655 GCACAGGATTTGCTGGCTTAGACGTAACTCTGCGGCTCAAGATGTCGCGCCACAGAGACAG 3714
3766 CGCTGGAGAACCGGGGAAACGAGCTGGGTACGCCCCAGAGACTGACCGGCTCTGGGCCAG 3825
3715 CGCTGGAGAACCGGGGAAACGAGCTGGGTACGCCCCAGAGACTGACCGGCTCTGGGCCAG 3774
3826 ACGCGTGGGCGGGTACAGGCTGGGCACTGACCCCGCAGCAGCTCTCAGGCGCTGCTTC 3885
3775 ACGCGTGGGCGGGTACAGGCTGGGCACTGACCCCGCAGCAGCTCTCAGGCGCTGCTTC 3834
3886 TCAAGCGCTTTCCTGCTTGGCGCGCAGCGCGCGCTGTTTCGCGCCAGATCGTGTGTC 3945

3835 TCAAGCGCTTTCCTGCTTGGCGCGCAGCGCGCGCTGTTTCGCGCCAGATCGTGTGTC 3894
3946 CTGCCCCCTTTGTGGGCTTGGGCTCTGTTTACGCTCTCATCTGCTCTCTTTTGGGCACT 4005
3895 CTGCCCCCTTTGTGGGCTTGGGCTCTGTTTACGCTCTCATCTGCTCTCTTTTGGGCACT 3954
4006 ACCCGGCTCTGGGCTCAGTCCACCATGTACGGTCTCAGGTGTCTTCTTTCAGTGAAG 4065
3955 ACCCGGCTCTGGGCTCAGTCCACCATGTACGGTCTCAGGTGTCTTCTTTCAGTGAAG 4014
4066 ACCCGGCTCTGGGCTCAGTCCACCATGTACGGTCTCAGGTGTCTTCTTTCAGTGAAG 4125
4015 ACCCGGCTCTGGGCTCAGTCCACCATGTACGGTCTCAGGTGTCTTCTTTCAGTGAAG 4074
4126 TGGAGAGAGCCCGAGTGGACATAGTCTCCACAGGTTCTTGGGACACAGAAATTCCTGCTG 4185
4075 TGGAGAGAGCCCGAGTGGACATAGTCTCCACAGGTTCTTGGGACACAGAAATTCCTGCTG 4134
4186 AAGTGGCCAAAGTCTTGGGCTCAGTGGCACTTGACCCACAGATCTCCATCCCAAGCTGCTCC 4245
4135 AAGTGGCCAAAGTCTTGGGCTCAGTGGCACTTGACCCACAGATCTCCATCCCAAGCTGCTCC 4194
4246 AAGTGTAGCAGCCCGTGGGCTGCTGCTGCCGACTGCCCGGCTGCAAGCTGGTGGTC 4305
4195 AAGTGTAGCAGCCCGTGGGCTGCTGCTGCCGACTGCCCGGCTGCAAGCTGGTGGTC 4254
4306 CCGCTCGGCGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTTCAAGAACTTGAAGCGGGA 4365
4255 CCGCTCGGCGCCAGGAGTGAACCGGCTCTGGGGAAGTGGTTTCAAGAACTTGAAGCGGGA 4314
4366 ACCTGTCTGACTTCTGCTCAAGACTACCGCGCTGTGGCCAGGCTGCAAGACTA 4425
4315 ACCTGTCTGACTTCTGCTCAAGACTACCGCGCTGTGGCCAGGCTGCTGCAAGACTA 4374
4426 AGAAGTGGGTGAATGAGGTACGAGTACGAGGCTTCTCGCTGGGGGCGCAGACCCAGGCTCC 4485
4375 AGAAGTGGGTGAATGAGGTACGAGTACGAGGCTTCTCGCTGGGGGCGCAGACCCAGGCTCC 4434
4486 TGCCCTCGGCGCCAGAGTGGGCGCTCAGTGGAGAGTGTGGGCGCTGCTGAGTCCCC 4545
4435 TGCCCTCGGCGCCAGAGTGGGCGCTCAGTGGAGAGTGTGGGCGCTGCTGAGTCCCC 4494
4546 TGCTCGGCGGCGCTCGACCGTCTGAAACCTCAGAGCTGGGCTCAGAGCTGG 4605
4495 TGCTCGGCGGCGCTCGACCGTCTGAAACCTCAGAGCTGGGCTCAGAGCTGG 4554
4606 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTGGCTT 4665
4555 ATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGGTGGCTT 4614
4666 TTGTCAACCGAGCAGCAACGCAATCTCGTGTCTCACTGCGCGCCAGGCGCGCGCGCT 4725
4615 TTGTCAACCGAGCAGCAACGCAATCTCGTGTCTCACTGCGCGCCAGGCGCGCGCGCT 4674
4726 AGCCCCACAGACTACCACTCAACCCCTTGAACCTCACCAGGAGCAGCTGTCTG 4785
4675 AGCCCCACAGACTACCACTCAACCCCTTGAACCTCACCAGGAGCAGCTGTCTG 4734
4786 AGGCTGCACTGATGGGCTCTCGTGGGAGCTCTCGTCTCCATCTGTGTGGTCTTTGGCA 4845
4735 AGGCTGCACTGATGGGCTCTCGTGGGAGCTCTCGTCTCCATCTGTGTGGTCTTTGGCA 4794
4846 TGTCTTTGTCCGCGCAGCTTCACTCTGTGTCTCATTTGAGGAGCAGTCAACCGAGCA 4905
4795 TGTCTTTGTCCGCGCAGCTTCACTCTGTGTCTCATTTGAGGAGCAGTCAACCGAGCA 4854
4906 AGCACCTGAGCTCATGGGCGCTGCTCCCACTCTACTGGCTTGGCACTTTCTCT 4965
4855 AGCACCTGAGCTCATGGGCGCTGCTCCCACTCTACTGGCTTGGCACTTTCTCT 4914
4966 GGCACATGTGTAACTATTGGTGCACAGCATGCTGTGTGCTCATCTTTCTGGGCTTCC 5025

D	b	4915	GGGACATGTATACTACTTGGTGCCAGCATGCATCGTGGTGCTCACTTTTCGGCCTTCC	4974
Q	y	5026	AGCAGAGGCCATATGTGGCCCCCTGCAAACCTGCCTGCTCTCCTGCTGTGTGCTACTACTGT	5085
D	b	4975	AGCAGAGGCCATATGTGGCCCCCTGCCAAACCTGCCTGCTCTCCTGCTGTGTGCTACTACTGT	5034
Q	y	5086	ATGGCTGCTCATCACACCGCTCATGTACCAGCGCTCTTCTTCTTCTTCGTGCCCAGCA	5145
D	b	5035	ATGGCTGCTCATCACACCGCTCATGTACCAGCGCTCTTCTTCTTCTTCGTGCCCAGCA	5094
Q	y	5146	CAGCCTATGTGCTCATCCTGCGATAAACCTCTTATTGGGCATCAATGGAGAGCATGCCA	5205
D	b	5095	CAGCCTATGTGCTCATCCTGCGATAAACCTCTTATTGGGCATCAATGGAGAGCATGCCA	5154
Q	y	5206	CCTTTGTCTTGAGCTCTTCTCATAGAAAGCTGCAGAGAGTGAGCCGATCTTTGAAC	5265
D	b	5155	CCTTTGTCTTGAGCTCTTCTCATAGAAAGCTGCAGAGAGTGAGCCGATCTTTGAAC	5214
Q	y	5266	AGGTCTTCTTATCTTCCCCACACTTCTGCTTTGGCCCGGGGGCTCATTTGAACATGTGCGGA	5325
D	b	5215	AGGTCTTCTTATCTTCCCCACACTTCTGCTTTGGCCCGGGGGCTCATTTGAACATGTGCGGA	5274
Q	y	5326	ACCAGGCCATGCTGATGSCCTTTTCAGCGCTTTGGGAGACAGCGAGTCCAGTCAACCCCTGC	5385
D	b	5275	ACCAGGCCATGCTGATGSCCTTTTCAGCGCTTTGGGAGACAGCGAGTCCAGTCAACCCCTGC	5334
Q	y	5386	GCTGGGAGTGGTCCGCAAGAACCTCTTTGGCCATGTGTATACAGGGGCCCCCTCTTCCCTTC	5445
D	b	5335	GCTGGGAGTGGTCCGCAAGNAACCTCTTTGGCCATGTGTATACAGGGGCCCCCTCTTCCCTTC	5394
Q	y	5446	TCTTTCACACTACTGCTGAGCAACCGAAGCCAACTCTGTCACACAGCCAGGGTGAGTCTC	5505
D	b	5395	TCTTTCACACTACTGCTGAGCAACCGAAGCCAACTCTGTCACACAGCCAGGGTGAGTCTC	5454
Q	y	5506	TGCCACTCTCGGAGAGAGGACAGCAGATGTAGCCCGTGAAACGGGAGCGGGTGGTCCAAG	5565
D	b	5455	TGCCACTCTCGGAGAGAGGACAGCAGATGTAGCCCGTGAAACGGGAGCGGGTGGTCCAAG	5514
Q	y	5566	GAGCCACCCAGGGGGATGTGTGGTGCTGAGGAACTTTGACCAAGGTATACCGTGGGCGAGA	5625
D	b	5515	GAGCCACCCAGGGGGATGTGTGGTGCTGAGGAACTTTGACCAAGGTATACCGTGGGCGAGA	5574
Q	y	5626	GGATGCCAGCTGTTGAACCGCTTGTGCTCGGGGATTCGCCCTGGTGAGTGTTTTGGGCTGC	5685
D	b	5575	GGATGCCAGCTGTTGAACCGCTTGTGCTCGGGGATTCGCCCTGGTGAGTGTTTTGGGCTGC	5634
Q	y	5686	TGGGTGTGAATGGAGCAGGGAAGAGCTCCAGTTTCCATGTGTGACCGGGGACACATTGG	5745
D	b	5635	TGGGTGTGAATGGAGCAGGGAAGAGCTCCAGTTTCCATGTGTGACCGGGGACACATTGG	5694
Q	y	5746	CCAGCAGGGCGAGGCTGTGCTGCGAGGCCACAGCGTGGCCCGGGAACCCAGTGTGCGC	5805
D	b	5695	CCAGCAGGGCGAGGCTGTGCTGCGAGGCCACAGCGTGGCCCGGGAACCCAGTGTGCGC	5754
Q	y	5806	ACCTCAGCATGGGATCTGCCCTCAATCCGATGCCATCTTTTGAGTGTCTGACGGGCGCG	5865
D	b	5755	ACCTCAGCATGGGATCTGCCCTCAATCCGATGCCATCTTTTGAGTGTCTGACGGGCGCG	5814
Q	y	5866	AGCACTGTGAGCTCTTGGCGGCTGCGCGGTGTCCCGAGAGGCCAGGTTGCCAGACCG	5925
D	b	5815	AGCACTGTGAGCTCTTGGCGGCTGCGCGGTGTCCCGAGAGGCCAGGTTGCCAGACCG	5874
Q	y	5926	CTGGCTCAGGCCCTGGCGGCTCTGGAGCTCTCATGTATCGCAGACCGGCCCTGCAGGCACCT	5985
D	b	5875	CTGGCTCAGGCCCTGGCGGCTCTGGAGCTCTCATGTATCGCAGACCGGCCCTGCAGGCACCT	5934
Q	y	5986	ACAGCGGAGGGAAACAAACGCAAGCTGCGCAGCGGCCCTTGGCGCTGTGTGGGGAACCCAGCCG	6045
D	b	5935	ACAGCGGAGGGAAACAAACGCAAGCTGCGCAGCGGCCCTTGGCGCTGTGTGGGGAACCCAGCCG	5994
Q	y	6046	TGTTGTTTTGACAGCGCACACAGGCATGGAACCCAGCGCGCGCGCTCTCTTTGGA	6105
D	b	5995	TGTTGTTTTGACAGCGCACACAGGCATGGAACCCAGCGCGCGCGCTCTCTTTGGA	6054

RESULT 4

US-10-154-419-1

US-10-134-413-1
: Sequence 1, Application US/10154419

; Publication No. US20030143675A1

; GENERAL INFORMATION:

; APPLICANT: Curtis, Rory A.J.

APPLICANT: Glucksman, Maria Alexandra

APPLICANT: Mevers, Rachel E

; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,

TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,

67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
TITLE OF INVENTION:

; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-249

; CURRENT APPLICATION NUMBER: US/10/154,419

; CURRENT FILING DATE: 2002-05-22

; Prior Application remove

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: Fa

; SEQ ID NO 1

; LENGTH: 6

; TYPE: DNA

; ORGANISM: Homo sapiens

[illegible]

4306 CCCTCCGCCCCGAGTGACCGGCTCTGGGAAAGTGGTTCAAGAACTGACAGGCGGA 4365
4255 CCCTCCGCCCCGAGTGACCGGCTCTGGGAAAGTGGTTCAAGAACTGACAGGCGGA 4314
4366 ACCTGTCTGACTTCTCTGGTCAAGACTTAACCGGCGCTGGTGGCGGAGGCTGAAGACTA 4425
4315 ACCTGTCTGACTTCTCTGGTCAAGACTTAACCGGCGCTGGTGGCGGAGGCTGAAGACTA 4374
4426 AGAAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 4485
4375 AGAAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 4434
4486 TGCCCTCGGCGCAAGAGTTCGGCGGCTCACTGAGGAGGAGTGGCGGCTGCTGAGTCCCC 4545
4435 TGCCCTCGGCGCAAGAGTTCGGCGGCTCACTGAGGAGGAGTGGCGGCTGCTGAGTCCCC 4494
4546 TGCTGCGCGGGCCCTCGACCGTGTCTGAAACCTTCAAGAGCTGGGCTCAAGAGCTGG 4605
4495 TGCTGCGCGGGCCCTCGACCGTGTCTGAAACCTTCAAGAGCTGGGCTCAAGAGCTGG 4554
4606 ATGCTCAGGACAGTCTCAAGACTGAGTTCAGAAAGCTTCAAGAGCTGGGCTCAAGAGCTGG 4665
4555 ATGCTCAGGACAGTCTCAAGACTGAGTTCAGAAAGCTTCAAGAGCTGGGCTCAAGAGCTGG 4614
4666 TTGTCACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCGCCCGGCGGCGGCGGCGG 4725
4615 TTGTCACCGAGCCAGCAACGCAATCTCGTGTCTCACTGCGCCCGGCGGCGGCGGCGGCGG 4674
4726 ACGCCACAGGATCACCACACTCAACACCGCTTGAACCTTCAACAGGAGCAGCTGTCTG 4785
4675 ACGCCACAGGATCACCACACTCAACACCGCTTGAACCTTCAACAGGAGCAGCTGTCTG 4734
4786 AGGCTGCACTGATGGCCCTCTCGGTGGAAGTCTCGTGTCTCACTGAGTGGTCTTGGCCA 4845
4735 AGGCTGCACTGATGGCCCTCTCGGTGGAAGTCTCGTGTCTCACTGAGTGGTCTTGGCCA 4794
4846 TGTCTTTGTCGGCGGAGCTTCACTTGTCTGCTTCACTGAGGAGCAGTCAACCGAGCCA 4905
4795 TGTCTTTGTCGGCGGAGCTTCACTTGTCTGCTTCACTGAGGAGCAGTCAACCGAGCCA 4854
4906 AGCACCTGCAAGCTCATGGGGGGCTGTCCCGCCACCTCTACTGGCTTGGCAACTTCTCT 4965
4855 AGCACCTGCAAGCTCATGGGGGGCTGTCCCGCCACCTCTACTGGCTTGGCAACTTCTCT 4914
4966 GGGACATGTTGTAATACTTGGTGCAGCATGCAATGGTGGTGTCTCACTTTTGGCCCTTCC 5025
4915 GGGACATGTTGTAATACTTGGTGCAGCATGCAATGGTGGTGTCTCACTTTTGGCCCTTCC 4974
5026 AGCAGAGGATATGTTGGCCCTTCCCAACTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 5085
4975 AGCAGAGGATATGTTGGCCCTTCCCAACTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 5034
5086 ATGGCTGGTGCATCACAACCGCTCATGTACCGAGCTCTTCTTCTTCTTCTTCTTCTTCTT 5145
5035 ATGGCTGGTGCATCACAACCGCTCATGTACCGAGCTCTTCTTCTTCTTCTTCTTCTTCTT 5094
5146 CAGCCTATGTTGGTGTCACTGATCAAGAACTCTTTATTTGATCAATGGAAGCATGGCCA 5205
5095 CAGCCTATGTTGGTGTCACTGATCAAGAACTCTTTATTTGATCAATGGAAGCATGGCCA 5154
5206 CTTTGTGCTGAGCTTCTCTGATCAGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5265
5155 CTTTGTGCTGAGCTTCTCTGATCAGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5214
5266 AGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5325
5215 AGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5274
5326 ACCAGGCAATGGCTGATGCTTTTGGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5385
5275 ACCAGGCAATGGCTGATGCTTTTGGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5334
5386 GCTGGGAGGTTGGTGGGCAAGAACTCTTGGGCAATGGTGTGATACAGGGGCGGCTCTTCTTCTT 5445

5335 GCTGGGAGGTTGGTGGGCAAGAACTCTTGGCCATGATGATACAGGGGCGGCTCTTCTTCTT 5394
5446 TCTTCACTAATCTGCTGAGACACGAAAGCAACTCTCTGCCACAGCCAGGAGGAGGAGGAGG 5505
5395 TCTTCACTAATCTGCTGAGACACGAAAGCAACTCTCTGCCACAGCCAGGAGGAGGAGGAGG 5454
5506 TGCCATCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5565
5455 TGCCATCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5514
5566 GAGCCACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5625
5515 GAGCCACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5574
5626 GAGTGCAGCTGTTGACCCGCTTGTCTGGGAGTCCCTCTGTGTGAGTGTGTTTGGGCTGC 5685
5575 GAGTGCAGCTGTTGACCCGCTTGTCTGGGAGTCCCTCTGTGTGAGTGTGTTTGGGCTGC 5634
5686 TGGGTGTGATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5745
5635 TGGGTGTGATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5694
5746 CCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5805
5695 CCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5754
5806 ACCTCAGCATGGGATCTGCCCTCAATCCGATCCATCTTTGAGCTGTCTGACGGGCGGCG 5865
5755 ACCTCAGCATGGGATCTGCCCTCAATCCGATCCATCTTTGAGCTGTCTGACGGGCGGCG 5814
5866 AGCACCTGAGAGCTTGTGGCGCTTGGCGGCTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5925
5815 AGCACCTGAGAGCTTGTGGCGCTTGGCGGCTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5874
5926 CTGGCTCAGGCTTGGCGCTTGGGAGCTCTCACTGTGACGAGACCCGCGCTTCCAGGAGCACT 5985
5875 CTGGCTCAGGCTTGGCGCTTGGGAGCTCTCACTGTGACGAGACCCGCGCTTCCAGGAGCACT 5934
5986 ACAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6045
5935 ACAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5994
6046 TGGTGTCTTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6105
5995 TGGTGTCTTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6054
6106 ACAGCCTTTTGGCGCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6165
6055 ACAGCCTTTTGGCGCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6114
6166 AGAGGTGTGAAGCGCTCTGCTCGCGCTTACCACTCATGGTGAATGGGCGGTTCCGCTGCC 6225
6115 AGAGGTGTGAAGCGCTCTGCTCGCGCTTACCACTCATGGTGAATGGGCGGTTCCGCTGCC 6174
6226 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6285
6175 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6234
6286 TGCCCGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6345
6235 TGCCCGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6294
6346 AGCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6405
6295 AGCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6354
6406 CCCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6465
6355 CCCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6414
6466 TTTTCCGTGAGCCAGAGAGTGTGGAGGAGGAGTATTTTGTACTTCTTCAAGGAGCAGGGGGA 6525

2374	ACACGCTGATGTGGGGCACTCTGTGTGGGCA	CGCTTGGGCGGAGTACGAGAGTGCCTGTCTT	2433
1436	GGACAAGCTGGAGCGGCACCTCAGAGGCAGCC	CTGGTGTCTGCGGGGCCCTCGCAACTGCCT	1495
2434	GGACNAGCTGGAGCGGCACCTCAGAGGCAGCC	CTGGTGTCTGCGGGGCCCTCGCAACTGCCT	2493
1496	CGCGGAACATCGATTCTTGGGCGGGCGTCTT	CTTGGGACCTGAGGACTCTTTCAGACCC	1555
2494	CGCGGAACATCGATTCTTGGGCGGGCGTCTT	CTTGGGACTCTTTCAGACCC	2553
1556	CACAGACGCCAACCCACGACCTGGGCCCGGCC	ACGTGGCGCATCAAAATCCGCATGGA	1615
2554	CACAGACGCCAACCCACGACCTGGGCCCGGCC	ACGTGGCGCATCAAAATCCGCATGGA	2613
1616	CATTGACGTGTCACGAGGACCAATAGATCAGG	GACAGGCTTTTGGGACCTCGGCCAGC	1675
2614	CATTGACGTGTCACGAGGACCAATAGATCAGG	GACAGGCTTTTGGGACCTCGGCCAGC	2673
1676	CGCGGACCCCTGACCGACTCTGGCTACGTGTG	GGGGCGGCTTCGTGTACTCTGCAAGACCT	1735
2674	CGCGGACCCCTGACCGACTCTGGCTACGTGTG	GGGGCGGCTTCGTGTACTCTGCAAGACCT	2733
1736	GGTGGAGCTGACGCGTCTGCGGTCTCAGGGCG	CCAAACCCCGGGCGGCTCTACCT	1795
2734	GGTGGAGCTGACGCGTCTGCGGTCTCAGGGCG	CCAAACCCCGGGCGGCTCTACCT	2793
1796	GCAGCAGATGCCCTATCCGTGCTATGTGGAC	AGAGTGTCTCTCGTGTCTGAGCCGGTC	1855
2794	GCAGCAGATGCCCTATCCGTGCTATGTGGAC	AGAGTGTCTCTCGTGTCTGAGCCGGTC	2853
1856	GCTGCCCTCTTCTTGACGCTGGCCTGGATCT	ACTCTCGTGACACTGACAGTGAAGCCGT	1915
2854	GCTGCCCTCTTCTTGACGCTGGCCTGGATCT	ACTCTCGTGACACTGACAGTGAAGCCGT	2913
1916	GGTGGGGAGAGGAGACGGGCTGCGGGACACA	CCATGCGCGCCATGGGGCTCAGCCGCGC	1975
2914	GGTGGGGAGAGGAGACGGGCTGCGGGACACA	CCATGCGCGCCATGGGGCTCAGCCGCGC	2973
1976	GGTGCTCTGGCTAGGCTGTCTCAGCTGCCTC	GGGCCCTTCTCTGCTCAGCGCCGGCT	2035
2974	GGTGCTCTGGCTAGGCTGTCTCAGCTGCCTC	GGGCCCTTCTCTGCTCAGCGCCGGCT	3033
2036	GCTGGTTCTGGTGCTCAAGCTGGGGGACATC	CTCCCTACAGCCACCCGGCGTGTCTT	2095
3034	GCTGGTTCTGGTGCTCAAGCTGGGGGACATC	CTCCCTACAGCCACCCGGCGTGTCTT	3093
2096	CCTGTTCTTGGCAGCCTTCGGGTGGGCCAC	GGTGACCCAGAGCTTCTGTCTCAGCGCTT	2155
3094	CCTGTTCTTGGCAGCCTTCGGGTGGGCCAC	GGTGAGCCAGAGCTTCTGTCTCAGCGCTT	3153
2156	CTTCTCCCGCGCAACCTGCTCGGGCTCGGG	CGGCTGGGCTACTTCTCCCTCTACCT	2215
3154	CTTCTCCCGCGCAACCTGCTCGGGCTCGGG	CGGCTGGGCTACTTCTCCCTCTACCT	3213
2216	GCCTTACGTCTGTGTGTGCTTTGGCGGAC	CGGCTGCGGGTGGCCGCGTGGCCGC	2275
3214	GCCTTACGTCTGTGTGTGCTTTGGCGGAC	CGGCTGCGGGTGGCCGCGTGGCCGC	3273
2276	GAGCCTGCTGTGCGCCGTGGCTTCGGCTTC	GGCTGGGAGAGCCTGGCTCTGTGGAGGA	2335
3274	GAGCCTGCTGTGCGCCGTGGCTTCGGCTTC	GGCTGGGAGAGCCTGGCTCTGTGGAGGA	3333
2336	GCAGGGGAGGGCGCAGTGGACACAGTGGG	CACCCGGCTTACGGCAGACGCTTTCAG	2395
3334	GCAGGGGAGGGCGCAGTGGACACAGTGGG	CACCCGGCTTACGGCAGACGCTTTCAG	3393
2396	CCTGGCCCAAGTCTCTCGGCCCTTCTGTCT	GCTGGAAGCGCGGCTCTACGGCCCTGCACCTG	2455
3394	CCTGGCCCAAGTCTCTCGGCCCTTCTGTCT	GGAAGCGCGGCTCTACGGCCCTGCACCTG	3453
2456	GTACTCTGGAAAGCTGTGTGCCAGGCCAGTA	CGGGATCCCTGAACCATGGAAATTTCTCTT	2515
3454	GTACTCTGGAAAGCTGTGTGCCAGGCCAGTA	CGGGATCCCTGAACCATGGAAATTTCTCTT	3513

Qy	2516	TCGGAGGAGCTACTGTGTGGGA	CCTCGGCCCCCAAGAGTCA	GAGCCCTTTCGCCCA	CCCC	2575
Db	3514	TCGGAGGAGCTACTGTGTGGGA	CCTCGGCCCCCAAGAGTCA	GAGCCCTTTCGCCCA	CCCC	3573
Qy	2576	GCTGGACCCAAAGTCTGTAGAA	GAGGACCGCCGCTCAGTCA	GGCTTCTGGCCTATCCGT	2635	
Db	3574	GCTGGACCCAAAGTCTGTAGAA	GAGGACCGCCGCTCAGTCA	GGCTTCTGGCCTATCCGT	3633	
Qy	2636	TCGCAGCTCGGAGAAGCGCTTTC	CTCGAAGCCCGCAGCCCTCG	GGGGCTCAGCCT	2695	
Db	3634	TCGCAGCTCGGAGAAGCGCTTTC	CTCGAAGCCCGCAGCCCTCG	GGGGCTCAGCCT	3693	
Qy	2696	GGACTTCTACAGGGCCACATCA	ACGCGCTTCTTGGGCAAA	CGGGCCCGCAAGACCAC	2755	
Db	3694	GGACTTCTACAGGGCCACATCA	ACGCGCTTCTTGGGCAAA	CGGGCCCGCAAGACCAC	3753	
Qy	2756	CACCTGTCCATCTTGAGTGGCC	CTTCCACCCAGTGGTCTGCT	GCCTTCATCTCTGGG	2815	
Db	3754	CACCTGTCCATCTTGAGTGGCC	CTTCCACCCAGTGGTCTGCT	GCCTTCATCTCTGGG	3813	
Qy	2816	CCACGACGTCCGCTCCAGCATG	CGCGCCATCCCGGCCACCTGG	GGCGTCTGTCTCAGTA	2875	
Db	3814	CCACGACGTCCGCTCCAGCATG	CGCGCCATCCCGGCCACCTGG	GGCGTCTGTCTCAGTA	3873	
Qy	2876	CAACGTGCTGTTTGACATGTGA	ACGTGAGACGACAGTCTGTG	TTCTATGGGCGGCTGAA	2935	
Db	3874	CAACGTGCTGTTTGACATGTGA	ACGTGAGACGACAGTCTGTG	TTCTATGGGCGGCTGAA	3933	
Qy	2936	GGGTCTGAGTGC	CGCTGTAGTGGCCCGCAGCAG	GCCTCTCTCAGAGATCTGGGCT	2995	
Db	3934	GGGTCTGAGTGC	CGCTGTAGTGGCCCGCAGCAG	GCCTCTCTCAGAGATCTGGGCT	3993	
Qy	2996	GGTCTCAAGCAGAGTGTG	CAGACTCGCCACCTCTCTGG	TGGGATCAACGGAACTGTC	3055	
Db	3994	GGTCTCAAGCAGAGTGTG	CAGACTCGCCACCTCTCTGG	TGGGATCAACGGAACTGTC	4053	
Qy	3056	CGTGGCCATTGCTTTGTGGG	CGGCTCCCAAGTTGTTATCT	TGGACAGGCTACGGCTGG	3115	
Db	4054	CGTGGCCATTGCTTTGTGGG	CGGCTCCCAAGTTGTTATCT	TGGACAGGCTACGGCTGG	4113	
Qy	3116	CGTGGATCTCTGCTTCCGCG	CGGTATTTGGGAGCTGCTCT	CAAAATCCCGAAAGGTCTG	3175	
Db	4114	CGTGGATCTCTGCTTCCGCG	CGGTATTTGGGAGCTGCTCT	CAAAATCCCGAAAGGTCTG	4173	
Qy	3176	CACGCTGATCCTCTCCAC	CCACCACTCTGATAGG	CAGAGCTGCTGGAGACCGTGTGC	3235	
Db	4174	CACGCTGATCCTCTCCAC	CCACCACTCTGATAGG	CAGAGCTGCTGGAGACCGTGTGC	4233	
Qy	3236	TGTGTTGGCAGTGGCGCTTGT	GCTGTGCTCCCACTCTCTCT	CTCGCGCGCTCACCT	3295	
Db	4234	CGTGGTGGCAGTGGCGCTTGT	GCTGTGCTCCCACTCTCTCT	CTCGCGCGCTCACCT	4293	
Qy	3296	GGGCTCCGGCTACTACTG	ACGCTGTGAAGGCCGCTTGC	CCCTCAGCACCAACATGAGAA	3355	
Db	4294	GGGCTCCGGCTACTACTG	ACGCTGTGAAGGCCGCTTGC	CCCTCAGCACCAACATGAGAA	4353	
Qy	3356	GGCTGACACTGACATG	AGGGCAGTGTGACAC	CAGGCAGGAAAGAAATGGCAGCCA	3415	
Db	4354	GGCTGACACTGACATG	AGGGCAGTGTGACAC	CAGGCAGGAAAGAAATGGCAGCCA	4413	
Qy	3416	GGGCAGCAGATCGGCATCTCT	CAGCTGTGCGCTTACGACT	TGGGTGCCGGGC	3475	
Db	4414	GGGCAGCAGATCGGCATCTCT	CAGCTGTGCGCTTACGACT	TGGGTGCCGGGC	4473	
Qy	3476	ACGGCTGGTGGAGAGCTG	CCACA	CGAGCTGGTGTCTGCTCCCTTAC	CGGGTGCCCA	3535
Db	4474	ACGGCTGGTGGAGAGCTG	CCACA	CGAGCTGGTGTCTGCTCCCTTAC	CGGGTGCCCA	4533
Qy	3536	TGACGGCAGCTTTCGCCA	CACTCTTCGAGAGCTAG	ACACGGCTGGCGAGCTGAGCT	3595	
Db	4534	TGACGGCAGCTTTCGCCA	CACTCTTCGAGAGCTAG	ACACGGCTGGCGAGCTGAGCT	4593	

3596 QY CACTGGCTACGGGATCTCGGACACAGCCCTCAGGAGATCTTCTGAAAGGTGGTGAGGA 3655
4594 Db CACTGGCTACGGGATCTCGGACACAGCCCTCAGGAGATCTTCTGAAAGGTGGTGAGGA 4653
3656 QY GTGTGCTCGGACACAGATATGGAGGATGGCAGCTGCGGGCAGCACCTATGACACAGGCAT 3715
4654 Db GTGTGCTCGGACACAGATATGGAGGATGGCAGCTGCGGGCAGCACCTATGACACAGGCAT 4713
3716 QY TGCTGGCCCTAGACGTAAACCTCTCGGCTCAAGATGCGGCCACAGGAGACAGCGCTCGAGAA 3775
4714 Db TGCTGGCCCTAGACGTAAACCTCTCGGCTCAAGATGCGGCCACAGGAGACAGCGCTCGAGAA 4773
3776 QY CGGGAAACAGCTGGGCTAGCCCAAGAGACTGACAGAGGCTCTGGGCGAGAGCGCGTGGG 3835
4774 Db CGGGAAACAGCTGGGCTAGCCCAAGAGACTGACAGAGGCTCTGGGCGAGAGCGCGTGGG 4833
3836 QY CGGGTACAGGGCTGGGCACTGACCCGACAGCCTGACAGCTCCAGGCGCTTCTCAAGCGCTT 3895
4834 Db CGGGTACAGGGCTGGGCACTGACCCGACAGCCTGACAGCTCCAGGCGCTTCTCAAGCGCTT 4893
3896 QY TGTGCTTGGCCGCGCAGCGCGCGGCTGTTTCGCCAGATGCTGCTGCCCTCTT 3955
4894 Db TGTGCTTGGCCGCGCAGCGCGCGGCTGTTTCGCCAGATGCTGCTGCCCTCTT 4953
3956 QY TGTGGGCTGGCCCTCGTGTTCAGCCTCATCTGTGCTCTCTTTTGGGCACTACCGGCTCT 4015
4954 Db TGTGGGCTGGCCCTCGTGTTCAGCCTCATCTGTGCTCTCTTTTGGGCACTACCGGCTCT 5013
4016 QY GCGGCTCAGTCCACCATGTAAGGTGCTCAGGTGCTTCTTCAGTGAGGACGCCCCAGG 4075
5014 Db GCGGCTCAGTCCACCATGTAAGGTGCTCAGGTGCTTCTTCAGTGAGGACGCCCCAGG 5073
4076 QY GGACCTTGGAAGTGCCTCGGCTGCTCGAGCGCTGCTGCGAGGAGCAGACTGGAGAGCC 4135
5074 Db GGAACCTTGGAAGTGCCTCGGCTGCTCGAGCGCTGCTGCGAGGAGCAGACTGGAGAGCC 5133
4136 QY CCCAGTGCAGCATAGCTCCACAGGTTCTCGCACACAGAGTTCCTGCTGAAGTGGCCAA 4195
5134 Db CCCAGTGCAGCATAGCTCCACAGGTTCTCGCACACAGAGTTCCTGCTGAAGTGGCCAA 5193
4196 QY GGTCTTGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCTGCTGAGTGGCCAA 4255
5194 Db GGTCTTGCCAGTGGCACTGGACCCAGAGTCTCCATCCCGAGCTGCTGCTGAGTGGCCAA 5253
4256 QY GCGCGGCTCCCGGCGCTGCTGCCGAGCTGCCCGGCTGAGCTGGTGGTCCCGCTCCGCC 4315
5254 Db GCGCGGCTCCCGGCGCTGCTGCCGAGCTGCCCGGCTGAGCTGGTGGTCCCGCTCCGCC 5313
4316 QY CCAGCAGTGAACCGCTCTGGGGAAGTGGTTCAAGAACCTGACAGCCCGGAACCTGTCTGA 4375
5314 Db CCAGCAGTGAACCGCTCTGGGGAAGTGGTTCAAGAACCTGACAGCCCGGAACCTGTCTGA 5373
4376 QY CTTCTGTGTCAAGACTTACCGCGCTGGTGGCCAGGCGCTGAAGACTAAGAGTGGGT 4435
5374 Db CTTCTGTGTCAAGACTTACCGCGCTGGTGGCCAGGCGCTGAAGACTAAGAGTGGGT 5433
4436 QY GAA TGAGGTCAAGTACGAGGCTTCTCGCTGGGGGCGAGACCCAGGCGCTGCCCTGGG 4495
5434 Db GAA TGAGGTCAAGTACGAGGCTTCTCGCTGGGGGCGAGACCCAGGCGCTGCCCTGGG 5493
4496 QY CCAAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCGCTGCCCTGGCGG 4555
5494 Db CCAAGAGTGGGCGCTCAGTGGAGGAGTGTGGGCGCTGCTGAGTCCCGCTGCCCTGGCGG 5553
4556 QY GCGCCTCGACGCTGTCTGAAACCTCAACAGCTGGGCTCAGAGCTGGATGCTCAGGA 4615
5554 Db GCGCCTCGACGCTGTCTGAAACCTCAACAGCTGGGCTCAGAGCTGGATGCTCAGGA 5613
4616 QY CAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTTGGCTTTTGTCAACCG 4675
5614 Db CAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGGCACTCCATGTTGGCTTTTGTCAACCG 5673
4676 QY AGCCAGCAACGCAATCTCTCGTGTCTCACTGCCCCAGGCGCGGCCCGCCACGCCCACAG 4735

5674 Db AGCCAGCAACGCAATCTCTCGTGTCTCACTGCCCCAGGCGCGGCCCGCCACAG 5733
4736 QY CATCACACACTCAACCAACCCCTTGAACTCACCAAGGAGCAGCTGCTGAGGCTGCAT 4795
5734 Db CATCACACACTCAACCAACCCCTTGAACTCACCAAGGAGCAGCTGCTGAGGCTGCAT 5793
4796 QY GATGCGCTCTCGGTGGACGCTCTCGTCTCATCTGTGTGGTCTTTTGCCATGCTCTTTGT 4855
5794 Db GATGCGCTCTCGGTGGACGCTCTCGTCTCATCTGTGTGGTCTTTTGCCATGCTCTTTGT 5853
4856 QY CCGGCGCAGCTTCACTCTTGTCTCATTTGAGGAGGAGTCAACCCAGGCAAGCACTGCA 4915
5854 Db CCGGCGCAGCTTCACTCTTGTCTCATTTGAGGAGGAGTCAACCCAGGCAAGCACTGCA 5913
4916 QY GCTCATGGGGGCGCTGTCCCGACCCCTCTACTGGCTTGGCACTTCTCTGGGACATGTG 4975
5914 Db GCTCATGGGGGCGCTGTCCCGACCCCTCTACTGGCTTGGCACTTCTCTGGGACATGTG 5973
4976 QY TAACTACTTGGTGGCAGCATGCTGCTGCTCATCTTTCTGGGCTTCCAGCAGAGGCG 5035
5974 Db TAACTACTTGGTGGCAGCATGCTGCTGCTCATCTTTCTGGGCTTCCAGCAGAGGCG 6033
5036 QY ATATGTGGCCCTGCGCAACCTGCTCTCTGCTGTGTCTACTGTATGCTGTGCTG 5095
6034 Db ATATGTGGCCCTGCGCAACCTGCTCTCTGCTGTGTCTACTGTATGCTGTGCTG 6093
5096 QY GATCACACCGCTCATGTACCCAGCCTCTTCTTCTCGTGCCGACGACAGCCTATGT 5155
6094 Db GATCACACCGCTCATGTACCCAGCCTCTTCTTCTCGTGCCGACGACAGCCTATGT 6153
5156 QY GGTGCTCACCTGCGATAAACCTCTTTATTGGCATCAATGGAAGCATGGCCACCTTTGTCT 5215
6154 Db GGTGCTCACCTGCGATAAACCTCTTTATTGGCATCAATGGAAGCATGGCCACCTTTGTCT 6213
5216 QY TGAGCTCTTCTCTGATCAGAACTGCAAGAGTGAAGCGGATCTTTGAAACAGGTCTTCT 5275
6214 Db TGAGCTCTTCTCTGATCAGAACTGCAAGAGTGAAGCGGATCTTTGAAACAGGTCTTCTCT 6273
5276 QY TATCTTCCCCACATCTGCTTGGGCGCGGCTCATTTGACATGCTGGGACCCAGGCGCAT 5335
6274 Db TATCTTCCCCACATCTGCTTGGGCGCGGCTCATTTGACATGCTGGGACCCAGGCGCAT 6333
5336 QY GGCTGATGCTCTTGGAGCGCTTGGGAGCAGGAGTTCAGTCAACCTGCGCTGGAGGT 5395
6334 Db GGCTGATGCTCTTGGAGCGCTTGGGAGCAGGAGTTCAGTCAACCTGCGCTGGAGGT 6393
5396 QY GGTGGCAAGAACCTCTTGGCATGGTGATACAGGGGCGCTCTTCTTCTTCTTCACT 5455
6394 Db GGTGGCAAGAACCTCTTGGCATGGTGATACAGGGGCGCTCTTCTTCTTCTTCACT 6453
5456 QY ACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAGCCAGGCTGAGGTCTTGCCTCTCT 5515
6454 Db ACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAGCCAGGCTGAGGTCTTGCCTCTCT 6513
5516 QY GGGAGAGGAGCAGAGATGTAGCCCGTGAAACGGAGCGGCTGGTCCAAAGAGCACCACA 5575
6514 Db GGGAGAGGAGCAGAGATGTAGCCCGTGAAACGGAGCGGCTGGTCCAAAGAGCACCACA 6573
5576 QY GGGGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5635
6574 Db GGGGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6633
5636 QY TGTGTGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5695
6634 Db TGTGTGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6693
5696 QY TGGAGCAGGGAGAGCTCCAGCTTTCGCACTGGTGAACGGGGACACATTTGGGCGAGGCGG 5755
6694 Db TGGAGCAGGGAGAGCTCCAGCTTTCGCACTGGTGAACGGGGACACATTTGGGCGAGGCGG 6753
5756 QY CGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5815

Db 721 AGCTTCGGGCACTGCTGCAGAGACCCCGAGGACACGCGGCCCTCGAGTTGCTGTTCAG 780
Qy 874 AGGCCCTCTGAGTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACG 933
Db 781 AGGCCCTCTGAGTGTTCAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACG 840
Qy 934 AGGCTAGTGAACCTGANTGAGTGTGTGGGCAAGGACCCAGAAATCCGCTTGCACAGACA 993
Db 841 AGGCTAGTGAACCTGANTGAGTGTGTGGGCAAGGACCCAGAAATCCGCTTGCACAGACA 900
Qy 994 GCCTGAGCCCGCTGCTCGAGCTGTGATTGAGCCCTTGGACAGCCCGCTGTCCCGCC 1053
Db 901 GCCTGAGCCCGCTGCTCGAGCTGTGATTGAGCCCTTGGACAGCCCGCTGTCCCGCC 960
Qy 1054 TGCTCTGAGACCGCTGAAGCTCTGATCTCTCGGAAGCTACTCTTTTGACACAGATACAC 1113
Db 961 TGCTCTGAGACCGCTGAAGCTCTGATCTCTCGGAAGCTACTCTTTTGACACAGATACAC 1020
Qy 1114 CTTTAAACCCGAAGCTCATGSCCCCAAGTGAAACCGGACCTTCGAGGAGCTCACCTGTCTGA 1173
Db 1021 CTTTAAACCCGAAGCTCATGSCCCCAAGTGAAACCGGACCTTCGAGGAGCTCACCTGTCTGA 1080
Qy 1174 GGGATGTCGGGAGGTGTGGAGATGCTTGGACCCCGGATCTTCACTTCATGAACGACA 1233
Db 1081 GGGATGTCGGGAGGTGTGGAGATGCTTGGACCCCGGATCTTCACTTCATGAACGACA 1140
Qy 1234 GTTCCAAATGTGGCCATGCTGCAGCGGCTCTGCAAGATCAGAGTGAAGGAAGGAGGACG 1293
Db 1141 GTTCCAAATGTGGCCATGCTGCAGCGGCTCTGCAAGATCAGAGTGAAGGAAGGAGGACG 1200
Qy 1294 CCAGACCTGGAGGCGGGACCAATGAGAGGCCCTGCGATCCCTTTCTGAGCCCTGGGAGCG 1353
Db 1201 CCAGACCTGGAGGCGGGACCAATGAGAGGCCCTGCGATCCCTTTCTGAGCCCTGGGAGCG 1260
Qy 1354 GTGGCTACAGTGGAGGACCGACACGCTGATGTGGGCACTCGTGGGACGCTGGGCC 1413
Db 1261 GTGGCTACAGTGGAGGACCGACACGCTGATGTGGGCACTCGTGGGACGCTGGGCC 1320
Qy 1414 GAGTGAAGGAGTGTCTGCTTGGACAGCTGAGGAGGCGGACCCCTCAGAGGACGCTTGG 1473
Db 1321 GAGTGAAGGAGTGTCTGCTTGGACAGCTGAGGAGGCGGACCCCTCAGAGGACGCTTGG 1380
Qy 1474 TGTCCGGGCCCTGCAATCTGCTCGGGAAATCGATTTCTGGGCGGGGCTCGTCTTCTTGG 1533
Db 1381 TGTCCGGGCCCTGCAATCTGCTCGGGAAATCGATTTCTGGGCGGGGCTCGTCTTCTTGG 1440
Qy 1534 GACCTGAGGAATCTTTTCAACCCCAAGAGCACCCAAACCCCAAGACCTGGGCCCGGCAACG 1593
Db 1441 GACCTGAGGAATCTTTTCAACCCCAAGAGCACCCAAACCCCAAGACCTGGGCCCGGCAACG 1500
Qy 1594 TGCGCATCAAAATCCGCAATGACATTTGACGTTGATGAGGACCAATGAATCAGGGACA 1653
Db 1501 TGCGCATCAAAATCCGCAATGACATTTGACGTTGATGAGGACCAATGAATCAGGGACA 1560
Qy 1654 GGTTTTGGGACCTTGGCCCAAGCGGACCCCTGACCGACCTGCGCTACGTTGAGGCGG 1713
Db 1561 GGTTTTGGGACCTTGGCCCAAGCGGACCCCTGACCGACCTGCGCTACGTTGAGGCGG 1620
Qy 1714 GCCTTGTGTACCTGCAAGA CTTGTGTGAGCGTGAAGCGGCTCCGCGTCTCAGCGCGCA 1773
Db 1621 GCCTTGTGTACCTGCAAGA CTTGTGTGAGCGTGAAGCGGCTCCGCGTCTCAGCGCGCA 1680
Qy 1774 ACCCCGGGGCGGCTCTTACCTGACGAGATGCCCTATCCGTGATGTGAACACGCTGT 1833
Db 1681 ACCCCGGGGCGGCTCTTACCTGACGAGATGCCCTATCCGTGATGTGAACACGCTGT 1740
Qy 1834 TCCTGCGTGTCTGAGCGGCTGCTGCGGCTCTTCTGAGCGCTGGCTGATCTACTCCG 1893
Db 1741 TCCTGCGTGTCTGAGCGGCTGCTGCGGCTCTTCTGAGCGCTGGCTGATCTACTCCG 1800
Qy 1894 TGAACACTGACAGTGAAGCCCGTGTGTGGGAGAGGAGACGCGGCTGCGGACACCATGC 1953

Db 1801 TGAACACTGACAGTGAAGCCCGTGTGTGGGAGAGGAGACGCGGCTGCGGAGACACCATGC 1860
Qy 1954 CGCCCATGGGCTCAGCCGCGGCTGTCTCTGGCTAGGCTGGTTCTCAGCTGCTCTCGGAC 2013
Db 1861 CGCCCATGGGCTCAGCCGCGGCTGTCTCTGGCTAGGCTGGTTCTCAGCTGCTCTCGGAC 1920
Qy 2014 CTTTCTGCTCAGCGCCGCGCTGTGTGTGTGTGTCTCAAGCTGGGAGACATCTCTCCCT 2073
Db 1921 CTTTCTGCTCAGCGCCGCACTGTGTGTGTGTGTGTCTCAAGCTGGGAGACATCTCTCCCT 1980
Qy 2074 ACAGCCACCCGGGCGTGTCTTCTTCTTGGCAGCCTTCCGCTGGGAGACGCTGTGACCC 2133
Db 1981 ACAGCCACCCGGGCGTGTCTTCTTCTTGGCAGCCTTCCGCTGGGAGACGCTGTGACCC 2040
Qy 2134 AGAGCTTCTCTCAGCGGCTTCTTCTTCCCGGCAACCTTGGCTGGGCTGCGGCGGCC 2193
Db 2041 AGAGCTTCTCTCAGCGGCTTCTTCTTCCCGGCAACCTTGGCTGGGCTGCGGCGGCC 2100
Qy 2194 TGGCCTACTTCTCTCTTACCTGCTTACCTGCTGTGTGTGTGTGTGGGAGACGCTGTG 2253
Db 2101 TGGCCTACTTCTCTCTTACCTGCTTACCTGCTGTGTGTGTGTGTGGGAGACGCTGTG 2160
Qy 2254 CCGGGGTGGCCGCTGGCCGCGAGCCTGCTGTGCTGCTGCTGCTGCTTGGCTTGGCTGG 2313
Db 2161 CCGGGGTGGCCGCTGGCCGCGAGCCTGCTGTGCTGCTGCTGCTTGGCTTGGCTGG 2220
Qy 2314 AGAGCTTGGCTCTCTGAGGAGAGGCGGAGGCGGCGCAGTGGCAACCTGGGACCC 2373
Db 2221 AGAGCTTGGCTCTCTGAGGAGAGGCGGAGGCGGCGCAGTGGCAACCTGGGACCC 2280
Qy 2374 GGCCTACGCGAGACGCTTTCAGCCTGCGCCCAAGGCTCTTGGGCTTCTGTGTGTGACGCG 2433
Db 2281 GGCCTACGCGAGACGCTTTCAGCCTGCGCCCAAGGCTCTTGGGCTTCTGTGTGTGACGCG 2340
Qy 2434 CGCTCTAGCGGCTCGCCACCTGGTACCTGAGAGCTGTGTGCTGAGGCGGACGCTACCGGATCC 2493
Db 2341 CGCTCTAGCGGCTCGCCACCTGGTACCTGAGAGCTGTGTGCTGAGGCGGACGCTACCGGATCC 2400
Qy 2494 CTGAAACATATGAAATTTTCTTTTCGAGGAGCTACTGTGTGGAGCTCTCGGCCCCCAAGA 2553
Db 2401 CTGAAACATATGAAATTTTCTTTTCGAGGAGCTACTGTGTGGAGCTCTCGGCCCCCAAGA 2460
Qy 2554 GTCCAGCCCTTGGCCCAACCCCGCTGGA CCAAGGCTGTGTGTGAGAGGACGCGCCG 2613
Db 2461 GTCCAGCCCTTGGCCCAACCCCGCTGGA CCAAGGCTGTGTGTGAGAGGACGCGCCG 2520
Qy 2614 GCCTGAGTCTGGGCTATCCGTTGCGAGCTGGAAGAGCGCTTCTCGGAAAGCGGACG 2673
Db 2521 GCCTGAGTCTGGGCTATCCGTTGCGAGCTTCCGCTTCCGAGAGCGCTTCTCGGAAAGCGGACG 2580
Qy 2674 CAGCCCTGCGGGGCTCAGCCTGACCTTAC CAGGCGCACATCACGCTTCTCTGGGCC 2733
Db 2581 CAGCCCTGCGGGGCTCAGCCTGACCTTAC CAGGCGCACATCACGCTTCTCTGGGCC 2640
Qy 2734 ACAAACGCGGCGGCAAGACCA CCACTGTCCATCTTGTAGTGGGCTCTTCCACCCAGTG 2793
Db 2641 ACAAACGCGGCGGCAAGACCA CCACTGTCCATCTTGTAGTGGGCTCTTCCACCCAGTG 2700
Qy 2794 GTGGCTCTGCTTCACTCTGGGCA CCAAGCTGCGCTCCAGCATGGCGGCTCTCGGCGCC 2853
Db 2701 GTGGCTCTGCTTCACTCTGGGCA CCAAGCTGCGCTCCAGCATGGCGGCTCTCGGCGCC 2760
Qy 2854 ACCTGGGCTGTCTCCTCAGTACA CAACTGTCTGTTTGA CATGCTGACCTGAGCAGGACG 2913
Db 2761 ACCTGGGCTGTCTCCTCAGTACA CAACTGTCTGTTTGA CATGCTGACCTGAGCAGGACG 2820
Qy 2914 TCTGTTCTATGGGCGGCTGAAAGGCTGTAGTGGCTGTAGTGGGCGGCGGAGGAC 2973
Db 2821 TCTGTTCTATGGGCGGCTGAAAGGCTGTAGTGGCTGTAGTGGGCGGCGGAGGAC 2880
Qy 2974 GTCTGTGAGGAGTGTGGGCTGTCTCCAAGCAGAGTGTGTCAGACTGCGCACCTCTCTG 3033
Db 2881 GTCTGTGAGGAGTGTGGGCTGTCTCCAAGCAGAGTGTGTCAGACTGCGCACCTCTCTG 2940

QY	3034	GTGGGATGCAACGGAAAGCTGTCCGTGGCCATTGCTTTGTGGGGCGCTCCCAAGTTGTTA	3093
Db	2941	GTGGGATGCAACGGAAAGCTGTCCGTGGCCATTGCTTTGTGGGGCGCTCCCAAGTTGTTA	3000
QY	3094	TCCTGGACGAGCCTACCGCTGGCGGTGGATCCTGCTTCCCGCCCGGTAATTTGGGAGCTGC	3153
Db	3001	TCCTGGACGAGCCTACCGCTGGCGGTGGATCCTGCTTCCCGCCCGGTAATTTGGGAGCTGC	3060
QY	3154	TGCTCAAAATACCGAAGAGTGCACCGCTGATTCCTCTCCACCCACCACTGGATGAGGCGAG	3213
Db	3061	TGCTCAAAATACCGAAGAGTGCACCGCTGATTCCTCTCCACCCACCACTGGATGAGGCGAG	3120
QY	3214	AGCTGCTGGAGACCGGTGGCTGTGGTGGACGGTGGCGCTGTGCTGCTGTGGCTCCC	3273
Db	3121	AGCTGCTGGAGACCGGTGGCTGTGGTGGACGGTGGCGCTGTGCTGCTGTGGCTCCC	3180
QY	3274	CACCTTTCTGCGCCGCTCACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGCC	3333
Db	3181	CACCTTTCTGCGCCGCTCACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGCC	3240
QY	3334	TGCCCTTGACCAACAATGAGAAAGGCTGACACTGACATGAGAGGCGAGTGTGGAACCAAGGC	3393
Db	3241	TGCCCTTGACCAACAATGAGAAAGGCTGACACTGACATGAGAGGCGAGTGTGGAACCAAGGC	3300
QY	3394	AGGAAAGAGATGGCAGCCAGGCGACGAGTCCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTG	3453
Db	3301	AGGAAAGAGATGGCAGCCAGGCGACGAGTCCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTG	3360
QY	3454	TACAGCACTGGCTGCCCGGGCCAGCTGGTGGAGGAGCTGCCACACGAGCTGGTGGTGG	3513
Db	3361	TACAGCACTGGCTGCCCGGGCCAGCTGGTGGAGGAGCTGCCACACGAGCTGGTGGTGG	3420
QY	3514	TGCTGCCCTTACA CGGGTGCCTATGACCGGACGCTTTCGCCACACTTTCCCGAGAGCTAGACA	3573
Db	3421	TGCTGCCCTTACA CGGGTGCCTATGACCGGACGCTTTCGCCACACTTTCCCGAGAGCTAGACA	3480
QY	3574	CGCGGCTGGCGGAGCTGAGCTCACTGGCTTACCGGATCTCCGACACGAGCTCGAGGGA	3633
Db	3481	CGCGGCTGGCGGAGCTGAGCTCACTGGCTTACCGGATCTCCGACACGAGCTCGAGGGA	3540
QY	3634	TCTTCTGAAAGTGGTGGAGAGTGTGCTCGGACACAGATATGGAGGATGGCAGCTGCG	3693
Db	3541	TCTTCTGAAAGTGGTGGAGAGTGTGCTCGGACACAGATATGGAGGATGGCAGCTGCG	3600
QY	3694	GGCAGCACCTATGCA CAGGCAATTTGCTGGCCCTAGACGTAACCCCTACGAGTGC	3753
Db	3601	GGCAGCACCTATGCA CAGGCAATTTGCTGGCCCTAGACGTAACCCCTACGAGTGC	3660
QY	3754	CACAGGACACGCTGGAGAA CGGGGAACACGCTGGGTGAGCCCGACGAGACTGACGAGG	3813
Db	3661	CACAGGACACGCTGGAGAA CGGGGAACACGCTGGGTGAGCCCGACGAGACTGACGAGG	3720
QY	3814	GCTCTGGGCCACAGCCCGTGGCGGGTACAGGGCTGGGCACTGACCCCGACGAGCTCC	3873
Db	3721	GCTCTGGGCCACAGCCCGTGGCGGGTACAGGGCTGGGCACTGACCCCGACGAGCTCC	3780
QY	3874	AGGCCCTGCTTCTCAAGCGCTTTCTGCTTCCCGCCGCGACGCCCGCGGCTGTTCCGCC	3933
Db	3781	AGGCCCTGCTTCTCAAGCGCTTTCTGCTTCCCGCCGCGACGCCCGCGGCTGTTCCGCC	3840
QY	3934	AGATGCTGCTGCTGCTGCTTGTGGGCTTGGCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTC	3993
Db	3841	AGATGCTGCTGCTGCTGCTTGTGGGCTTGGCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTC	3900
QY	3994	CTTTTGGGGCACTACCGGGCTCTGGGCTCAGTCCCACTATGATGATGATGATGATGATGATG	4053
Db	3901	CTTTTGGGGCACTACCGGGCTCTGGGCTCAGTCCCACTATGATGATGATGATGATGATGATG	3960
QY	4054	TCTTCAATGAGGACGCCCCAGGGGACCTTGGACGCTGCCCGGCTGCTGAGGCGCTGCTGC	4113
Db	3961	TCTTCAATGAGGACGCCCCAGGGGACCTTGGACGCTGCCCGGCTGCTGAGGCGCTGCTGC	4020

QY	4114	AGGAGCGAGACTGAGGAGGAGCCCGGCTGAGCATAGCTCCCA CAGGTTCTCGGCAACCAG	4173
Db	4021	AGGAGCGAGACTGAGGAGGAGCCCGGCTGAGCATAGCTCCCA CAGGTTCTCGGCAACCAG	4080
QY	4174	AGGTTCTCTGCTGAAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGA CCCCAGAGTCTCCAT	4233
Db	4081	AGGTTCTCTGCTGAAAGTGGCCAAAGTCTTGGCCAGTGGCAACTGGA CCCCAGAGTCTCCAT	4140
QY	4234	CCCCAGCTGCGCAGTGTAGCAGCCCGGCTGCGCGGCGCTGCTGCGCGGCTG	4293
Db	4141	CCCCAGCTGCGCAGTGTAGCAGCCCGGCTGCGCGGCGCTGCTGCGCGGCTG	4200
QY	4294	CAGCTGCTGGTCCCGCTCGCCCGCAGGCACTGACCGGCTCTGGGGAGTGGTTTCAGAACC	4353
Db	4201	CAGCTGCTGGTCCCGCTCGCCCGCAGGCACTGACCGGCTCTGGGGAGTGGTTTCAGAACC	4260
QY	4354	TGACAGGCGCGGAA CTTGTCTGACTTCTCTGGTCAAGACCTTACCCGCGCTGGTGC	4413
Db	4261	TGACAGGCGCGGAA CTTGTCTGACTTCTCTGGTCAAGACCTTACCCGCGCTGGTGC	4320
QY	4414	GCCTGAAAGACTAAGAAAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	4473
Db	4321	GCCTGAAAGACTAAGAAAGTGGTGAATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG	4380
QY	4474	GAGACCCAGGCTGCGCTCGGGCCAAAGATTGGGCGCTCAGTGGAGGAGTGTGGGCGC	4533
Db	4381	GAGACCCAGGCTGCGCTCGGGCCAAAGATTGGGCGCTCAGTGGAGGAGTGTGGGCGC	4440
QY	4534	TGCTGAGTCCCTGCTGCTGCGGGGCGCTCGA CCGTGTCTCTGAAAAA CTTCAACGCTGGG	4593
Db	4441	TGCTGAGTCCCTGCTGCTGCGGGGCGCTCGA CCGTGTCTCTGAAAAA CTTCAACGCTGGG	4500
QY	4594	CTCAGAGCTGATGCTCAGGACAGTCTCAGATCTGGTTCACACAAAGGCTGGCACT	4653
Db	4501	CTCAGAGCTGATGCTCAGGACAGTCTCAGATCTGGTTCACACAAAGGCTGGCACT	4560
QY	4654	CCATGCTGGCTTTTGTCAACCGAGCGAGCAACGCAATCTCGTGTCTCACTCCGCCCCAG	4713
Db	4561	CCATGCTGGCTTTTGTCAACCGAGCGAGCAACGCAATCTCGTGTCTCACTCCGCCCCAG	4620
QY	4714	GCCTGCGCCGCCACGCGCCACAGCATCACCACTCAACCA CCCCCTTGAACCTCACCAAGG	4773
Db	4621	GCCTGCGCCGCCACGCGCCACAGCATCACCACTCAACCA CCCCCTTGAACCTCACCAAGG	4680
QY	4774	AGCAGCTGTCTGAGGCTGACATGATGCGCTCTCGGTGAGACGCTCCTGCTCATCTGTG	4833
Db	4681	AGCAGCTGTCTGAGGCTGACATGATGCGCTCTCGGTGAGACGCTCCTGCTCATCTGTG	4740
QY	4834	TGCTCTTTGCGCATGTCTTTTGTCCCGGCGAGCTTCACTCTTGTCTCTCATTTGAGGAGCGAG	4893
Db	4741	TGCTCTTTGCGCATGTCTTTTGTCCCGGCGAGCTTCACTCTTGTCTCTCATTTGAGGAGCGAG	4800
QY	4894	TCA CCGGAGCGAAGCA CTTGCACTCATGCGGGGCGCTGTCCCGCCACCTCTACTGCGTTG	4953
Db	4801	TCA CCGGAGCGAAGCA CTTGCACTCATGCGGGGCGCTGTCCCGCCACCTCTACTGCGTTG	4860
QY	4954	GMACTTTTCTTGGGACATGTGTAACTATTGCTGCGAGCATGATGATGATGATGATGATGATG	5013
Db	4861	GMACTTTTCTTGGGACATGTGTAACTATTGCTGCGAGCATGATGATGATGATGATGATGATG	4920
QY	5014	TTCTGCGCTTCCAGCAGAGGCGCATATGCTGGCGCCCTCCCAACCTGCTGCTCTCTGCTGT	5073
Db	4921	TTCTGCGCTTCCAGCAGAGGCGCATATGCTGGCGCCCTCCCAACCTGCTGCTCTCTGCTGT	4980
QY	5074	TGCTACTACTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5133
Db	4981	TGCTACTACTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5040
QY	5134	CGGTGCCAGCAGCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5193
Db	5041	CGGTGCCAGCAGCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5100
QY	5194	GAAAGCATGGCCACTTTTGTGCTTGTGCTTCTCTGATCAGAAAGCTGCAGGAGGTGAGCC	5253

QY 274 CGGTCCAGCTCTGTGCGAATGCTGTGGCTCTCTTCTCTTCTTCTCATCTCTGTGGCTG 333
Db 181 CGGTCCAGCTCTGTGCGAATGCTGTGGCTCTCTTCTCTTCTTCTCATCTCTGTGGCTG 240
QY 334 TTCCGCCATCCCAACCGCCCTCGAGCACCATGAATGCCACTTCCCAAAACAAGCCACTGC 393
Db 241 TTCCGCCATCCCAACCGCCCTCGAGCACCATGAATGCCACTTCCCAAAACAAGCCACTGC 300
QY 394 CATCGGGGGCAGCGTGCCTGTGCTCCAGGGTCTCATCTGTAATGGAACAACACCTGCT 453
Db 301 CATCGGGGGCAGCGTGCCTGTGCTCCAGGGTCTCATCTGTAATGGAACAACACCTGCT 360
QY 454 TTCCGAGCTGACACCGGGGGAGAGCCCGGGCGCTTGAGCAACTTCAACGACTTCCCTGG 513
Db 361 TTCCGAGCTGACACCGGGGGAGAGCCCGGGCGCTTGAGCAACTTCAACGACTTCCCTGG 420
QY 514 TCTCCCGGCTGTAGCGGATGCCCGCACCTGTGTGGGAGGGCCAGTGCCCAAGGACGC 573
Db 421 TCTCCCGGCTGTAGCGGATGCCCGCACCTGTGTGGGAGGGCCAGTGCCCAAGGACGC 480
QY 574 TGGCTGGCTTAGGGAAGCTGATCGCACCGCTGAGGGGTGACACGAGCAGCGCCAGCTTC 633
Db 481 TGGCTGGCTTAGGGAAGCTGATCGCACCGCTGAGGGGTGACACGAGCAGCGCCAGCTTC 540
QY 634 AACCAACCAAGCAAGTCTCCACTGGAAACCAACCCATGCTGGATGTCGGGAGCTGCTGACGT 693
Db 541 AACCAACCAAGCAAGTCTCCACTGGAAACCAACCCATGCTGGATGTCGGGAGCTGCTGACGT 600
QY 694 CACTGTGTCGACCGGAATCCTCGGGTTGGCACTGGGCCCAAGGCCAGGAGCCCTTGCACA 753
Db 601 CACTGTGTCGCAACGAATCCTCGGGTTGGCACTGGGCCCAAGGCCAGGAGCCCTTGCACA 660
QY 754 GCTTGTGGAGGCGCTGAGGACCTGGCCAGAGCTCTGTGGCGCTGCGAGCCTGGTG 813
Db 661 GCTTGTGGAGGCGCTGAGGACCTGGCCAGAGCTCTGTGGCGCTGCGAGCCTGGTG 720
QY 814 AGCTTGGGCACTGCTGCAGAGACCCGAGGGACCAAGCCCGCTGGAGTGTGCTGCTAG 873
Db 721 AGCTTGGGCACTGCTGCAGAGACCCGAGGGACCAAGCCCGCTGGAGTGTGCTGCTAG 780
QY 874 AGGCCCTCTCAGTGTGAGGGACCTAGCACAGTGGGCCCTCCTCAACTGTGTACG 933
Db 781 AGGCCCTCTCAGTGTGAGGGACCTAGCACAGTGGGCCCTCCTCAACTGTGTACG 840
QY 934 AGGCTAGTGACCTGATGGAGCTGGTGGGCGAGGACAGAAATCCGCCCTGCCAGACGA 993
Db 841 AGGCTAGTGACCTGATGGAGCTGGTGGGCGAGGACAGAAATCCGCCCTGCCAGACGA 900
QY 994 GCCTGAGCCCGCTGCTCGGAGCTGATGGAGCCCTGGACAGCCACCCGCTGTCCGCG 1053
Db 901 GCCTGAGCCCGCTGCTCGGAGCTGATGGAGCCCTGGACAGCCACCCGCTGTCCGCG 960
QY 1054 TGCTCTGGAGACGCTGAAGCCTCTGATCCTCGGGAAGTACTCTTTGCAACAGATACAC 1113
Db 961 TGCTCTGGAGACGCTGAAGCCTCTGATCCTCGGGAAGTACTCTTTGCAACAGATACAC 1020
QY 1114 CTTTTACCCGGAACTCATGGCCAGGTGAAACCGGACCTTCGAGGAGCTCACCTGTGTA 1173
Db 1021 CTTTTACCCGGAACTCATGGCCAGGTGAAACCGGACCTTCGAGGAGCTCACCTGTGTA 1080
QY 1174 GGGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTATGAACGACA 1233
Db 1081 GGGATGTCGGGAGGTGTGGAGATGCTGGGACCCCGGATCTTCACTTATGAACGACA 1140
QY 1234 GTTCCAAATGTGGCAATGCTGCAGCGCTCCTGACAGATGACAGATGAAGAAAGGAGC 1293
Db 1141 GTTCCAAATGTGGCAATGCTGCAGCGCTCCTGACAGATGACAGATGAAGAAAGGAGC 1200
QY 1294 CCAGACTGGAGCGGGACCAATGAGAGCCCTGCGATCTTTCTGGACCTCGGAGCG 1353
Db 1201 CCAGACTGGAGCGGGACCAATGAGAGCCCTGCGATCTTTCTGGACCTCGGAGCG 1260
QY 1354 GTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGGACCTGTGGTGGGACCGTGGGCC 1413

Db 1261 GTGGCTACAGCTGGCAGAGACGACACGCTGATGTGGGGCACCTGTGGGACCGTGGGCC 1320
QY 1414 GAGTGACGAGTGCCTGTCTTTGGACAAGCTGGAGGGGACCCCTCAGAGGACAGCCCTGG 1473
Db 1321 GAGTGACGAGTGCCTGTCTTTGGACAAGCTGGAGGGGACCCCTCAGAGGACAGCCCTGG 1380
QY 1474 TGTTCGGGGCCCTGCAACTGCTCGCGGAACATCGATTCTGGGGCGGCTGCTCTCTGG 1533
Db 1381 TGTTCGGGGCCCTGCAACTGCTCGCGGAACATCGATTCTGGGGCGGCTGCTCTCTGG 1440
QY 1534 GACCTCAGGACTCTTTCAGAGACCCCAAGAGACCCCAACCCCAAGACCTTGGGCGGCCACG 1593
Db 1441 GACCTCAGGACTCTTTCAGAGACCCCAAGAGACCCCAACCCCAAGACCTTGGGCGGCCACG 1500
QY 1594 TGGCATCAAAATCCGCATGGACATTTGACGTGTGTACGAGGACCAATAGATCAGGACA 1653
Db 1501 TGGCATCAAAATCCGCATGGACATTTGACGTGTGTACGAGGACCAATAGATCAGGACA 1560
QY 1654 GGTTCGAGACCTTGGCCCGGACCCCTGACCGACCCCTGACCGTACGCTAGCTGTGGGG 1713
Db 1561 GGTTCGAGACCTTGGCCCGGACCCCTGACCGGACCCCTGACCGTACGCTGTGGGG 1620
QY 1714 GCTTCGTGTACCTGCAAGACCTTGGTGGAGCGTGACGCGTCCGCTGTCTCAGCGGCCA 1773
Db 1621 GCTTCGTGTACCTGCAAGACCTTGGTGGAGCGTGACGCGTCCGCTGTCTCAGCGGCCA 1680
QY 1774 ACCCCGGGCGGCTCTTACCTGACAGATGCCCTATCCGTGTATGTGGACGAGTGT 1833
Db 1681 ACCCCGGGCGGCTCTTACCTGACAGATGCCCTATCCGTGTATGTGGACGAGTGT 1740
QY 1834 TCTTCGTGTGTGAGCGCGTGCCTGCCCTCTTCTGAGCGTGGGCTGGATCTACTCG 1893
Db 1741 TCTTCGTGTGTGAGCGCGTGCCTGCCCTCTTCTGACGCTGGGCTGGATCTACTCG 1800
QY 1894 TGACACTGACAGTGAAGCGCTGGTGGGAGAAAGAGACGCGGCTGCGGACACCATGC 1953
Db 1801 TGACACTGACAGTGAAGCGCTGGTGGGAGAAAGAGACGCGGCTGCGGACACCATGC 1860
QY 1954 GGGCCATGGGGCTACGCGCGGGTGTCTGGCTAGGCTGGTTCCTCAGCTGCTCGGG 2013
Db 1861 GGGCCATGGGGCTACGCGCGGGTGTCTGGCTAGGCTGGTTCCTCAGCTGCTCGGG 1920
QY 2014 CTTTCCTGTGTGAGCGCGCTGCTGGTTCGTGTCTCAAGCTGGGGGACATCTCCCT 2073
Db 1921 CTTTCCTGTGTGAGCGCGCTGCTGGTTCGTGTCTCAAGCTGGGGGACATCTCCCT 1980
QY 2074 ACAGCCACCCGGGCTGTCTTCTTTCGACGCTTTCGCGTGGCCACCGGTGACCC 2133
Db 1981 ACAGCCACCCGGGCTGTCTTCTTTCGACGCTTTCGCGTGGCCACCGGTGACCC 2040
QY 2134 AGAGCTTCTGTCTACGCGCTTCTTCTCCGCGCCAACTGTGTGGGCTTGGGGGCG 2193
Db 2041 AGAGCTTCTGTCTACGCGCTTCTTCTCCGCGCCAACTGTGTGGGCTTGGGGGCG 2100
QY 2194 TGGCTTACTTCTCCCTTACCTGCGCTACGCTGTGTGTGGCTTGGCGGAGCGGCTGC 2253
Db 2101 TGGCTTACTTCTCCCTTACCTGCGCTACGCTGTGTGTGGCTTGGCGGAGCGGCTGC 2160
QY 2254 CCSCGGTGGCGCGGTGGCGGAGCTGTGTGCGCCGCTTTCGCTTTCGCTTTCGCTGCG 2313
Db 2161 CCSCGGTGGCGCGGTGGCGGAGCTGTGTGCGCCGCTTTCGCTTTCGCTTTCGCTGCG 2220
QY 2314 AGAGCTTGTCTGTGTGGAGGACAGGGCGCGAGTGGCAACAGTGGGACCC 2373
Db 2221 AGAGCTTGTCTGTGTGGAGGACAGGGCGCGAGTGGCAACAGTGGGACCC 2280
QY 2374 GGCCTACGGGACAGCTTTCAGCCTGGCCAGGTCTCTGGCCTTCTGCTGTGAGACGG 2433
Db 2281 GGCCTACGGGACAGCTTTCAGCCTGGCCAGGTCTCTGGCCTTCTGCTGTGAGACGG 2340
QY 2434 CGCTCTACGGCCTTGGCACCTGGTACTGGAGCTGTGTGCCCGAGGCAAGTACGGGATCC 2493

2341	CGCTCTACGGCGCTCGCCACCTCGTGTACCTTGGAAGCTGTGTGTCGCCACAGGCCAGTGATCGGGATCC	2400
QY	CTGAACCATGGAAATTTTCTTTTCGGAGGAGCTACTGTGTCGGACCTCTCGGCCCCCCCAAGA	2553
DB	CTGAACCATGGAAATTTTCTTTTCGGAGGAGCTACTGTGTCGGACCTCGGCCCCCCCAAGA	2460
QY	GTCAGAGCCCTTGCCCAACCCCGCTGGACCCAAAGGTGCTGTGTAGAAGAGCACCGCCCG	2613
DB	GTCAGAGCCCTTGCCCAACCCCGCTGGACCCAAAGGTGCTGTGTAGAAGAGCACCGCCCG	2520
QY	GCCTGAGTCTGGCGTATCCGTTGCGAGCGCTGGAGAACGCTTTCTCTGGAGGCCCGCAGC	2673
DB	GCCTGAGTCTGGCGTATCCGTTGCGAGCGCTGGAGAACGCTTTCTCTGGAGGCCCGCAGC	2580
QY	CAGCCCTCGGGGGCTCAGCCTTGCACTTTACACAGGGCCACATCACCGCCCTTCTCTGGGCC	2733
DB	CAGCCCTCGGGGGCTCAGCCTTGCACTTTACACAGGGCCACATCACCGCCCTTCTCTGGGCC	2640
QY	ACAACGGGGCCGGAAGACACACCTGTTCATCTTGAGTGGGCTTTTCCACCCAGTG	2793
DB	ACAACGGGGCCGGAAGACACACCTGTTCATCTTGAGTGGGCTTTTCCACCCAGTG	2700
QY	GTGGCTCTGCCCTTCACTCTGGGCCACAGCGTCCGCTCCAGCATGGCGCCCATCGGGCCCC	2853
DB	GTGGCTCTGCCCTTCACTCTGGGCCACAGCGTCCGCTCCAGCATGGCGCCCATCGGGCCCC	2760
QY	ACCTGGGGCGTCTGCTCCTCAGTACAACTGTGTTTGACATGCTGACCGTGGACGAGCAGC	2913
DB	ACCTGGGGCGTCTGCTCCTCAGTACAACTGTGTTTGACATGCTGACCGTGGACGAGCAGC	2820
QY	TCTGGTTCTATGGCGGGCTGAAGGTTCTGAGTGCCTGTGAGTGGGCCCCGAGCAGGACC	2973
DB	TCTGGTTCTATGGCGGGCTGAAGGTTCTGAGTGCCTGTGAGTGGGCCCCGAGCAGGACC	2880
QY	GTCCTGTCAGGATGTGGGGCTGCTCCTCAAGCAGAGTGTGCAGATCTGCCACCTCTCTG	3033
DB	GTCCTGTCAGGATGTGGGGCTGCTCCTCAAGCAGAGTGTGCAGATCTGCCACCTCTCTG	2940
QY	GTGGGATGCAACGGAAGCTGTCTGGTGGCCATTTGCTTTTGGGGCGGCTCCAAAGTTGTTA	3093
DB	GTGGGATGCAACGGAAGCTGTCTGGTGGCCATTTGCTTTTGGGGCGGCTCCAAAGTTGTTA	3000
QY	TCTTGGACGAGCTTACCGGCTGGCGTGAATCTGCTTCCCGCCGCGGTATTTGGAGAGCTGC	3153
DB	TCTTGGACGAGCTTACCGGCTGGCGTGAATCTGCTTCCCGCCGCGGTATTTGGAGAGCTGC	3060
QY	TGCTCAAAATACCGAAGAGGTGCGACGCTGATCTCTCTCACCCACCTGGATGAGGCAG	3213
DB	TGCTCAAAATACCGAAGAGGTGCGACGCTGATCTCTCTCACCCACCTGGATGAGGCAG	3120
QY	AGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTGCTGTGTGGCTTCCC	3273
DB	AGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGGCTTGTGCTGTGTGGCTTCCC	3180
QY	CACCTTTCTGCGCGTCACTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGCC	3333
DB	CACCTTTCTGCGCGTCACTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCCCGCC	3240
QY	TGCCCCGTGACCAACCAATAGAGGCTGACATGACATGGAGGGCAGTGTGGACACGAGC	3393
DB	TGCCCCGTGACCAACCAATAGAGGCTGACATGACATGGAGGGCAGTGTGGACACGAGC	3300
QY	AGGAAAAGAGAAATGGCAGCCAGGGCAGCAGATCGGCACCTCTCAGCTGCTGGCCCTGG	3453
DB	AGGAAAAGAGAAATGGCAGCCAGGGCAGCAGATCGGCACCTCTCAGCTGCTGGCCCTGG	3360
QY	TACAGCACTGGGTGCCCGGGGCAACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTGTGG	3513
DB	TACAGCACTGGGTGCCCGGGGCAACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTGTGG	3420
QY	TGCTGCCCTTACACGGGTGCCCATGACGGCAGCTTTCGCCACACTCTTCCGAGAGCTAGACA	3573
DB	TGCTGCCCTTACACGGGTGCCCATGACGGCAGCTTTCGCCACACTCTTCCGAGAGCTAGACA	3480

QY	3574	CGCGCTGGCGAGTGAGGCTCACTCGGCTACGGGATCTTCGGAACACAGGCTTCGAGAGA	3633
DB	3481	CGCGCTGGCGGAGCTGAGGCTCACTCGCTACGGGATCTCCGACACAGGCTTCGAGAGA	3540
QY	3634	TCCTCCCTGAAGTGTGAGAGAGTGCTCGGGACACAGATATGGAGATGGCAGCTGCG	3693
DB	3541	TCCTTCCTGAAGTGTGAGAGAGTGTCTCGGGACACAGATATGGAGATGGCAGCTGCG	3600
QY	3694	GGCAGCACCTATGACACAGGCATTGCTGGCCCTAGACGTAAACCTTCGGCTCAAGATGCCGC	3753
DB	3601	GGCAGCACCTATGACACAGGCATTGCTGGCCCTAGACGTAAACCTTCGGCTCAAGATGCCGC	3660
QY	3754	CACAGGAGACAGCGCTGAGAAACGGGGAAACAGCTGGGTAGCCCCAGAGACTGACAGG	3813
DB	3661	CACAGGAGACAGCGCTGAGAAACGGGGAAACAGCTGGGTAGCCCCAGAGACTGACAGG	3720
QY	3814	GCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTGGGCACGTACACCGCCACGACGTCC	3873
DB	3721	GCTCTGGGCGAGACGCGTGGGCGGGTACAGGGCTGGGCACGTACACCGCCACGACGTCC	3780
QY	3874	AGGCCCTGCTTCTCAAGCGCTTTCTGTCTGCCCGCCAGCGCGCGGCTGTGTGCC	3933
DB	3781	AGGCCCTGCTTCTCAAGCGCTTTCTGTCTGCCCGCCAGCGCGGCTGTGTGCC	3840
QY	3934	AGATCGTGTGCTGCCTCTTTGTGGGCTTGGCCCTCTGTGTTCAGCCTCATCGTGCTC	3993
DB	3841	AGATCGTGTGCTGCCTCTTTGTGGGCTTGGCCCTCTGTGTTCAGCCTCATCGTGCTC	3900
QY	3994	CTTTTGGGCACTACCGCGCTCTGGGCTCAGTCCCACCATTGACGTGCTCAGGTGTCCT	4053
DB	3901	CTTTTGGGCACTACCGCGCTCTGGGCTCAGTCCCACCATTGACGTGCTCAGGTGTCCT	3960
QY	4054	TCCTTCAGTGAGAGCGCCACAGGGGACCTCTGGAAGTGCCCGGCTGCTCGAGCGCTGCTGC	4113
DB	3961	TCCTTCAGTGAGAGCGCCACAGGGGACCTCTGGAAGTGCCCGGCTGCTCGAGCGCTGCTGC	4020
QY	4114	AGGAGGCAAGACTGGAGAGCCCCAGTGCAGCATAGCTCCCAAGAGTTCTCGGCACAG	4173
DB	4021	AGGAGGCAAGACTGGAGAGCCCCAGTGCAGCATAGCTCCCAAGAGTTCTCGGCACAG	4080
QY	4174	AAGTTCTGCTGAAGTGGCCAAAGTCTTTGGCGAGTGGCAAATGGAACCCAGAGTCTCCAT	4233
DB	4081	AAGTTCTGCTGAAGTGGCCAAAGTCTTTGGCGAGTGGCAAATGGAACCCAGAGTCTCCAT	4140
QY	4234	CCCCAGCCTGCGAGTGTAGCGCCGGTGCCTGGCGGCTGCTCGGACCTGCCCCGGCTG	4293
DB	4141	CCCCAGCCTGCGAGTGTAGCGCCGGTGCCTGGCGGCTGCTCGGACCTGCCCCGGCTG	4200
QY	4294	CAGCTGGTGGTCCCTCCCGCCCCAGGCACTGACCCGGCTCTGGGGAAGTGGTTCAAGAAC	4353
DB	4201	CAGCTGGTGGTCCCTCCCGCCCCAGGCACTGACCCGGCTCTGGGGAAGTGGTTCAAGAAC	4260
QY	4354	TGACAGGCGGGAAACCTGTCTGAATTCCTGTGTCAAGACTTACCCGCGCTGGTGGCCAGG	4413
DB	4261	TGACAGGCGGGAAACCTGTCTGAATTCCTGTGTCAAGACTTACCCGCGCTGGTGGCCAGG	4320
QY	4414	GCCTGAGACTTGAAGATGGGTGAATGAGGTACAGGTACGAGGCTTCTGCTGGGGGGCC	4473
DB	4321	GCCTGAGACTTGAAGATGGGTGAATGAGGTACAGGTACGAGGCTTCTGCTGGGGGGCC	4380
QY	4474	GAGACCCAGGCTCCCTCCGGGCCAAGATTGGGCCGCTCAGTGTGAGGAGTGTGTGGCGC	4533
DB	4381	GAGACCCAGGCTCCCTCCGGGCCAAGATTGGGCCGCTCAGTGTGAGGAGTGTGTGGCGC	4440
QY	4534	TGCTGAGTCCCTGCTGGCGGGGCCCTCGACCGTGTCTTGAAAAACCTCACAGCCTGGG	4593
DB	4441	TGCTGAGTCCCTGCTGGCGGGGCCCTCGACCGTGTCTTGAAAAACCTCACAGCCTGGG	4500
QY	4594	CTCACAGCTGGATGTCTCAGACAGTCTCAAGATCTGGGTTCAACAAAGGCTGGCACT	4653
DB	4501	CTCACAGCTGGACGCTCTCAGACAGTCTCAAGATCTGGGTTCAACAAAGGCTGGCACT	4560

QY	4654	CCATGGTGGCTTTGTCAACCGAGCCAGCAACGAATCTCCGTGCTCACCTGCCGCCAG	4713
Db	4561	CCATGGTGGCTTTGTCAACCGAGCCAGCAACGAATCTCCGTGCTCACCTGCCGCCAG	4620
QY	4714	GCCCGGCCCGCCACGCGCCACAGCATCACCACTCAACCAACCCCTTGAACCTCACCAAGG	4773
Db	4621	GCCCGGCCCGCCACGCGCCACAGCATCACCACTCAACCAACCCCTTGAACCTCACCAAGG	4680
QY	4774	AGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGAGCTCTCGTCTCATCTGTG	4833
Db	4681	AGCAGCTGTCTGAGGCTGCACTGATGGCTCTCTCGTGGAGCTCTCGTCTCATCTGTG	4740
QY	4834	TGCTCTTTGGCAGTCTCTTTGTCCTGGCCAGCTTCACTCTTGTCTCACTTGGAGGAGGAG	4893
Db	4741	TGCTCTTTGGCAGTCTCTTTGTCCTGGCCAGCTTCACTCTTGTCTCACTTGGAGGAGGAG	4800
QY	4894	TCACCGAGCAAGACCTTGACGCTCATGGGGGCGCTGTCCGCCACCCCTCTACTTGGCTTG	4953
Db	4801	TCACCGAGCAAGACCTTGACGCTCATGGGGGCGCTGTCCGCCACCCCTCTACTTGGCTTG	4860
QY	4954	GCAACTTTCTCTGGGACATGTGTAACTACTTGTGTGCAGCATGATGCTGTGCTCATCT	5013
Db	4861	GCAACTTTCTCTGGGACATGTGTAACTACTTGTGTGCAGCATGATGCTGTGCTCATCT	4920
QY	5014	TTCTGGCTTCCAGCAGAGGCAATATGTGGCCCTGCCACCTGCTCTCTCTCTCTCT	5073
Db	4921	TTCTGGCTTCCAGCAGAGGCAATATGTGGCCCTGCCACCTGCTCTCTCTCTCTCTCT	4980
QY	5074	TGCTACTTGTATGTGTGCTGATCACACCGCTCATGTACCCAGCGCTCTCTCTCTCTCT	5133
Db	4981	TGCTACTTGTATGTGTGCTGATCACACCGCTCATGTACCCAGCGCTCTCTCTCTCTCT	5040
QY	5134	CGTGCCACGACAGCCTATGTGGTGTCTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	5193
Db	5041	CGTGCCACGACAGCCTATGTGGTGTCTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	5100
QY	5194	GAAGCATGGCCACTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT	5253
Db	5101	GAAGCATGGCCACTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT	5160
QY	5254	GGATCTTTGAAACAGGCTCTTCTTATCTTCCGCCACTTCTGCTTGGGCGGGGCTCATTTG	5313
Db	5161	GGATCTTTGAAACAGGCTCTTCTTATCTTCCGCCACTTCTGCTTGGGCGGGGCTTATTTG	5220
QY	5314	ACATGTGCGGAACCGAGCCATGGCTGATGCTCTTTGAGCGCTTGGAGACAGGAGTTCC	5373
Db	5221	ACATGTGCGGAACCGAGCCATGGCTGATGCTCTTTGAGCGCTTGGAGACAGGAGTTCC	5280
QY	5374	AGTCACCCCTGCGCTGGAGGCTGGTGGCAAGAACCTCTTGGCCATGGTGATACAGGGGC	5433
Db	5281	AGTCACCCCTGCGCTGGAGGCTGGTGGCAAGAACCTCTTGGCCATGGTGATACAGGGGC	5340
QY	5434	CCCTCTTCTCTCTTTCACACTACTGTGTCAGACACCGAAGCCAACTCTTGCCACAGCCCA	5493
Db	5341	CCCTCTTCTCTCTTTCACACTACTGTGTCAGACACCGAAGCCAACTCTTGCCACAGCCCA	5400
QY	5494	GGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGCAGGAGATGTAGCCCGTGAACGGGAGC	5553
Db	5401	GGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGCAGGAGATGTAGCCCGTGAACGGGAGC	5460
QY	5554	GGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGCAGGAGATGTAGCCCGTGAACGGGAGC	5613
Db	5461	GGGTGAGGTCTCTGCCACTCTCTGGGAGAGGAGCAGGAGATGTAGCCCGTGAACGGGAGC	5520
QY	5614	ACCGTGGGACAGAGATGCCAGCTTGTGACCGCTTGTGCTTGGGATTTCCCTCTGGTGAGT	5673
Db	5521	ACCGTGGGACAGAGATGCCAGCTTGTGACCGCTTGTGCTTGGGATTTCCCTCTGGTGAGT	5580
QY	5674	GTTTTGGGCTGTCTGGGTGTAATGGAGCAGGGAGAGCTCCAGCTTTCGATGGTGAACGG	5733
Db	5581	GTTTTGGGCTGTCTGGGTGTAATGGAGCAGGGAGAGCTCCAGCTTTCGATGGTGAACGG	5640
QY	5734	GGGACACATTTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCGGGAAC	5793

RESULT 8

Db	5641	GGGACACATTTGGCCAGCAGGGCGAGGCTGTGCTGCGAGGCCACAGCGTGGCCCGGGAAC	5700
QY	5794	CCAGTCTCTGGCACCCTCAGCATGGGATACCTGCCCTCAATCCGATGCATCTTTGAGTGC	5853
Db	5701	CCAGTCTCTGGCACCCTCAGCATGGGATACCTGCCCTCAATCCGATGCATCTTTGAGTGC	5760
QY	5854	TGACGGGCGCGAGCACCCTGGAGCTGCTTGGCGCCTGCGCGTGTCCCGGAGCCCGAGG	5913
Db	5761	TGACGGGCGCGAGCACCCTGGAGCTGCTTGGCGCCTGCGCGTGTCCCGGAGCCCGAGG	5820
QY	5914	TTGCCCCAGACCCGCTGAGGCTGGCGCTTGGGACTCTCATGTGTACGACAGACCGGC	5973
Db	5821	TTGCCCCAGACCCGCTGAGGCTGGCGCTTGGGACTCTCATGTGTACGACAGACCGGC	5880
QY	5974	CTGACGGCACCTACAGCGAGGAAACAAACGAAAGCTGCGAGCGGCCCTGGCGCTGGTTG	6033
Db	5881	CTGACGGCACCTACAGCGAGGAAACAAACGAAAGCTGCGAGCGGCCCTGGCGCTGGTTG	5940
QY	6034	GGGACCCAGCGTGGTGTCTTGGACGAGCGGACACAGGACATGGACCCAGCGCGGC	6093
Db	5941	GGGACCCAGCGTGGTGTCTTGGACGAGCGGACACAGGACATGGACCCAGCGCGGC	6000
QY	6094	GCTTCTTTGGAAACAGCTTTTGGCGCTGGTGGGAGGCGGCTTCACTGATGCTCACCT	6153
Db	6001	GCTTCTTTGGAAACAGCTTTTGGCGCTGGTGGGAGGCGGCTTCACTGATGCTCACCT	6060
QY	6154	CCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTGCGCTAGCCATCATGTTGAATGGGC	6213
Db	6061	CCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTGCGCTAGCCATCATGTTGAATGGGC	6120
QY	6214	GGTTCCGCTGCTGGGACGCGGACATCTCAAGGCGAGATTCGCGGCGGGTCAACAC	6273
Db	6121	GGTTCCGCTGCTGGGACGCGGACATCTCAAGGCGAGATTCGCGGCGGGTCAACAC	6180
QY	6274	TGACCCCTGGGGTGGCGCGCAAGGTCAGCGCGGAGCGGCTTCTGTTGGCGGCGGAGT	6333
Db	6181	TGACCCCTGGGGTGGCGCGCAAGGTCAGCGCGGAGCGGCTTCTGTTGGCGGCGGAGT	6240
QY	6334	TCCCTGGGTGGAGCTGCGGAGGACATGGAGGCGGCTTCCAGCTGCGCGCGG	6393
Db	6241	TCCCTGGGTGGAGCTGCGGAGGACATGGAGGCGGCTTCCAGCTGCGCGCGG	6300
QY	6394	GAGGCGCTGCGGCTGGCGCGGCTTGGAGAGCTGCGGCTGACGCTGCGCGGAGCAGC	6453
Db	6301	GAGGCGCTGCGGCTGGCGCGGCTTGGAGAGCTGCGGCTGACGCTGCGCGGAGCAGC	6360
QY	6454	GCGTGGAGGACTTTTCCGTGAGCCAGACATGCTGGAGGAGGTATTTCTTGTACTTCTCCA	6513
Db	6361	GCGTGGAGGACTTTTCCGTGAGCCAGACATGCTGGAGGAGGTATTTCTTGTACTTCTCCA	6420
QY	6514	AGGACAGGGGAGGACAGGACACCGAAGAGCAGAGGAGGAGGAGTGGGAGTGAACC	6573
Db	6421	AGGACAGGGGAGGACAGGACACCGAAGAGCAGAGGAGGAGGAGTGGGAGTGAACC	6480
QY	6574	CCGCGCCAGGCTGACGACCCCAACCGCTGAGCCAGTTCTCGATGACCTTAGCACTG	6633
Db	6481	CCGCGCCAGGCTGACGACCCCAACCGCTGAGCCAGTTCTCGATGACCTTAGCACTG	6540
QY	6634	CCGAGACTGTGCTCTGAGCTCTCCCTCCCTGGGGGCGCGGGGAGGCGCTTGGGAATGGC	6693
Db	6541	CCGAGACTGTGCTCTGAGCTCTCCCTCCCTGGGGGCGCGGGGAGGCGCTTGGGAATGGC	6600
QY	6694	AAGGGCAAGGTAGAGTGCCTAGGAGCCCTGGAATCTCAGGCTGGCAGAGGGGCTGGTGCCT	6753
Db	6601	AAGGGCAAGGTAGAGTGCCTAGGAGCCCTGGAATCTCAGGCTGGCAGAGGGGCTGGTGCCT	6660
QY	6754	GGGAAAAATAAGAGAGGCTGGAGAGAACCCCTGCTGTGANA	6797
Db	6661	GGGAAAAATAAGAGAGGCTGGAGAGAACCCCTGCTGTGANA	6704

US-10-775-920-4
; Sequence 4, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Meigen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 6588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-4

Query Match 96.5%; Score 6565.6; DB 17; Length 6588;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6574; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy	210	ATGGCTTCTGGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCGGAGA	269
Db	1	ATGGCTTCTGGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCGGAGA	60
Qy	270	CAGCGGTCCAGCTCCTGGTGAAATGCTGTGGCCCTCTCTCTCTTCTTCTTCTTCTG	329
Db	61	CAGCGGTCCAGCTCCTGGTGAAATGCTGTGGCCCTCTCTCTCTTCTTCTTCTTCTG	120
Qy	330	GCTGTTCCCACTCCACCGCCCTGGAGACCAATGAATGCCATTTCCCAAAAGGCA	389
Db	121	GCTGTTCCCACTCCACCGCCCTGGAGACCAATGAATGCCATTTCCCAAAAGGCA	180
Qy	390	CTGCCATCGGGGGCACCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAAACC	449
Db	181	CTGCCATCGGGGGCACCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAAACC	240
Qy	450	TGCTTTCCGAGCTGACACCGGGGCGAGAGCCCGGGCCCTGAGCAATCAACGACTCC	509
Db	241	TGCTTTCCGAGCTGACACCGGGGCGAGAGCCCGGGCCCTGAGCAATCAACGACTCC	300
Qy	510	CTGCTCTCCGGCTGCTAGCCGATCCCGCACTGTGCTGGAGGGCCAGTCCCAACAG	569
Db	301	CTGCTCTCCGGCTGCTAGCCGATCCCGCACTGTGCTGGAGGGCCAGTCCCAACAG	360
Qy	570	ACGCTGGCTGGCTAGGGAAGCTGATCGCCAGCTGAGGGTGCACGACGACGCGCCAG	629
Db	361	ACGCTGGCTGGCTAGGGAAGCTGATCGCCAGCTGAGGGTGCACGACGACGCGCCAG	420
Qy	630	CCTCAACCAAACAGAGCTCTCCACTGGAACCAACCATGTGGATGTCCGGAGTGTCTG	689
Db	421	CCTCAACCAAACAGAGCTCTCCACTGGAACCAACCATGTGGATGTCCGGAGTGTCTG	480
Qy	690	ACGTCACTGCTGGCAGCGAATCCCTGGGGTGGCACTGGGCGCAAGCCAGAGCCCTTG	749
Db	481	ACGTCACTGCTGGCAGCGAATCCCTGGGGTGGCACTGGGCGCAAGCCAGAGCCCTTG	540
Qy	750	CACAGCTGTTGGAGGCGCTGAGGACCTGGCCAGAGAGCTCTTGGGCTGCGAGCCCTG	809
Db	541	CACAGCTGTTGGAGGCGCTGAGGACCTGGCCAGAGAGCTCTTGGGCTGCGAGCCCTG	600
Qy	810	GTGAGCTTCGGGCACTGCTGAGAGACCCCGAGGGACACGCGGCCCTCTGGAGTGTCTG	869
Db	601	GTGAGCTTCGGGCACTGCTGAGAGACCCCGAGGGACACGCGGCCCTCTGGAGTGTCTG	660
Qy	870	TCAGAGCCCTCTGCAAGTGTAGGGGACCTAGCAGCAGTGGGCCCCCTCCCTCAACTGG	929
Db	661	TCAGAGCCCTCTGCAAGTGTAGGGGACCTAGCAGCAGTGGGCCCCCTCCCTCAACTGG	720

Qy	930	TACGAGGCTAGTGACCTGATCGAGCTGTGTGGGCGAGGACCCAGAAATCCGCTTCGACAG	989
Db	721	TACGAGGCTAGTGACCTGATCGAGCTGTGTGGGCGAGGACCCAGAAATCCGCTTCGACAG	780
Qy	990	AGCAGCTTGAGCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCGCTGTCC	1049
Db	781	AGCAGCTTGAGCCCGCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCGCTGTCC	840
Qy	1050	CGCTGCTCTGGAGACGCTCTGATCTCTCGGGAAGCTACTCTTTGACACAGAT	1109
Db	841	CGCTGCTCTGGAGACGCTCTGATCTCTCGGGAAGCTACTCTTTGACACAGAT	900
Qy	1110	ACA CTTTATCCCGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGAGAGCTCACTCTG	1169
Db	901	ACA CTTTATCCCGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGAGAGCTCACTCTG	960
Qy	1170	CTGAGGAGTCTCGGAGGCTGTGGAGATGTGGGACCCCGGATCTTACCTTCTATGAAC	1229
Db	961	CTGAGGAGTCTCGGAGGCTGTGGAGATGTGGGACCCCGGATCTTACCTTCTATGAAC	1020
Qy	1230	GACAGTTCCAAATGTGGCCATCTGCAGCGGCTCTTCGACAGATGCGAGATGAAGGAAGAG	1289
Db	1021	GACAGTTCCAAATGTGGCCATCTGCAGCGGCTCTTCGACAGATGCGAGATGAAGGAAGAG	1080
Qy	1290	CAGCCAGACCTGGAGGCGGGACCAATGAGGCGCTTCGATCTCTTCTGACCCCTGGG	1349
Db	1081	CAGCCAGACCTGGAGGCGGGACCAATGAGGCGCTTCGATCTCTTCTGACCCCTGGG	1140
Qy	1350	AGCGGTGCTACAGTGGCAGGACGCAACACCTGATGTGGGGCACCTGGTGGGACAGCTG	1409
Db	1141	AGCGGTGCTACAGTGGCAGGACGCAACACCTGATGTGGGGCACCTGGTGGGACAGCTG	1200
Qy	1410	GGCGAGTGCAGGAGTGCCTGCTTGGACAAAGCTGGAGGGGCGCACCTCAGAGGACGCC	1469
Db	1201	GGCGAGTGCAGGAGTGCCTGCTTGGACAAAGCTGGAGGGGCGCACCTCAGAGGACGCC	1260
Qy	1470	CTGCTGTGCGGGGCTGCAACTGCTCGCGAAACATCGATTTGCGGCGGGCTCGCTTTC	1529
Db	1261	CTGCTGTGCGGGGCTGCAACTGCTCGCGAAACATCGATTTGCGGCGGGCTCGCTTTC	1320
Qy	1530	TTGGGACCTGAGGACTCTTCAGACCCACAGAGACCCCAACCCAGACCTTGGGCCCGGCG	1589
Db	1321	TTGGGACCTGAGGACTCTTCAGACCCCAACAGAGACCCCAACCCAGACCTTGGGCCCGGCG	1380
Qy	1590	CACGTGCGCATCAAAATCCGCAATTCACGCTGTCAGAGGACCAATAAGATCAGG	1649
Db	1381	CACGTGCGCATCAAAATCCGCAATTCACGCTGTCAGAGGACCAATAAGATCAGG	1440
Qy	1650	GACAGGTTTTGGGACCTTGGCCAGCGGACCCCTTGACCGACCTCGCTTACGCTGG	1709
Db	1441	GACAGGTTTTGGGACCTTGGCCAGCGGACCCCTTGACCGACCTCGCTTACGCTGG	1500
Qy	1710	GGCGGCTGCTGTACCTGCAAGA CTTGGTGAGGCTGAGCGCTCCGCTGCTCAGCGGC	1769
Db	1501	GGCGGCTGCTGTACCTGCAAGA CTTGGTGAGGCTGAGCGCTCCGCTGCTCAGCGGC	1560
Qy	1770	GCCAAACCCCGGCGGCTCTACCTGAGAGATGCCCTATCCGCTGCTATGTGAGACAC	1829
Db	1561	GCCAAACCCCGGCGGCTCTACCTGAGAGATGCCCTATCCGCTGCTATGTGAGACAC	1620
Qy	1830	GTGTTCTCTGCTGCTGAGCGGCTGCTGCGGCTCTTCTGACGCTGGCTGATCTTAC	1889
Db	1621	GTGTTCTCTGCTGCTGAGCGGCTGCTGCGGCTCTTCTGACGCTGGCTGATCTTAC	1680
Qy	1890	TCCGTGACATGACAGTGAAGGCGCTGTGGGAGAGAGAGAGACGCGGCTGCGGAGAC	1949
Db	1681	TCCGTGACATGACAGTGAAGGCGCTGTGGGAGAGAGAGAGACGCGGCTGCGGAGAC	1740
Qy	1950	ATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2009
Db	1741	ATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
Qy	2010	GGGCGCTTCTGCTCAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2069

1801 GGGCCCTTCCTGCTCAGCGCGCACTGCTGTTCTGCTCAAGCTGGAGACATCCTC 1860
2070 CCTACAGCACCGGGCGTGGTCTTCCTGTTCTTGGCAGCCCTTCGGGTGGCCACGGTG 2129
1861 CCTACAGCACCGGGCGTGGTCTTCCTGTTCTTGGCAGCCCTTCGGGTGGCCACGGTG 1920
2130 ACCAGAGCTTCCTGCTCAGGCGCTTCTTCCTCCGGCCCAACTGGCTGGGCTGGGCG 2189
1921 ACCAGAGCTTCCTGCTCAGGCGCTTCTTCCTCCGGCCCAACTGGCTGGGCGCTGGGCG 1980
2190 GGCCTGGCCCTACTTCTCCCTCTACTGCGCCCTAGCTGCTGTGTGTGGCTGGCGGACCGG 2249
1981 GGCCTGGCCCTACTTCTCCCTCTACTGCGCCCTAGCTGCTGTGTGTGGCTGGCGGACCGG 2040
2250 CTGCCCGCGGGTGGCGCGTGGCGGAGCTGCTGTGTCGCCCTGGCTTCGCTTCGGC 2309
2041 CTGCCCGCGGGTGGCGCGTGGCGGAGCTGCTGTGTCGCCCTGGCTTCGCTTCGGC 2100
2310 TGGAGAGCTTGCTGCTGCGAGGAGCAGGGCGAGGCGCGCACTGGCCACAAGCTGGGC 2369
2101 TGGAGAGCTTGCTGCTGCGAGGAGCAGGGCGAGGCGCGCACTGGCCACAAGCTGGGC 2160
2370 ACCGGGCTACGGCAGACGCTTTCAGCCTGGCCCGAGGCTCTGGCCTTCTGCTGCTGGAC 2429
2161 ACCGGGCTACGGCAGACGCTTTCAGCCTGGCCCGAGGCTCTGGCCTTCTGCTGCTGGAC 2220
2430 GCGGCGCTCTACGGCTCGCCACCTGGTACCTGGAAGCTGTGTGCCAGGCCAGTACGGG 2489
2221 GCGGCGCTCTACGGCTCGCCACCTGGTACCTGGAAGCTGTGTGCCAGGCCAGTACGGG 2280
2490 ATCCCTGAACCATGGAATTTTCTTTTCGAGGAGCTACTGCTGGAGCCCTCGGCCCCC 2549
2281 ATCCCTGAACCATGGAATTTTCTTTTCGAGGAGCTACTGCTGGAGCCCTCGGCCCCC 2340
2550 AAGAGTCCAGCCCTTGCCCCACCCCGCTGGACCCAAAGGTGTGTTAGAAAGGACACCG 2609
2341 AAGAGTCCAGCCCTTGCCCCACCCCGCTGGA CCCAAAGGTGTGTTAGAAAGGACACCG 2400
2610 CCGGCGCTGAGTCTGCGGTATCCGTTTCGAGCTGGAGAGCGCTTTCCTGGAAGCCCG 2669
2401 CCGGCGCTGAGTCTGCGGTCTCGTTTCGAGCTGGAGAGCGCTTTCCTGGAAGCCCG 2460
2670 CAGCGACCTTCGGGGGCTCAGGCTGGACTTCTACAGGGCCACATCA CGGCTTCCTG 2729
2461 CAGCGACCTTCGGGGGCTCAGGCTGGACTTCTACAGGGCCACATCA CGGCTTCCTG 2520
2730 GGCACAA CGGGGCGGCAAGACCAACACCTGTCCATCTTGAAGTGGCCCTTTC CCAACCC 2789
2521 GGCACAA CGGGGCGGCAAGACCAACACCTGTCCATCTTGAAGTGGCCCTTTC CCAACCC 2580
2790 AGTGGTGGCTCTGCTTCATCTGGGCGCA CGAGCTCGCTCCAGCATGGCGGCATCCGG 2849
2581 AGTGGTGGCTCTGCTTCATCTGGGCGCA CGAGCTCGCTCCAGCATGGCGGCATCCGG 2640
2850 CCCACCTTCGGGCTGCTCTCAGTACAA CGTGTGTTTGA CATGCTGA CGGTGGA CGAG 2909
2641 CCCACCTTCGGGCTGCTCTCAGTACAA CGTGTGTTTGA CATGCTGA CGGTGGA CGAG 2700
2910 CAGTCTGTGTTCTATGGGCGGCTGAAGGGCTGAGTGCCTGTGAGTGGGCCCGGAGCAG 2969
2701 CAGTCTGTGTTCTATGGGCGGCTGAAGGGCTGAGTGCCTGTGAGTGGGCCCGGAGCAG 2760
2970 GACGCTCTGTCAGGATGTGGGGCTGGTCTCCAAGCAGAGTGTGCA GACTCGGCCACCTC 3029
2761 GACGCTCTGTCAGGATGTGGGGCTGGTCTCCAAGCAGAGTGTGCA GACTCGGCCACCTC 2820
3030 TCTGTTGGGATGCA CGGAAGCTGTCCGTGGGCCATTGCTTTGTGGGCGGCTCCCAAGTT 3089
2821 TCTGTTGGGATGCA CGGAAGCTGTCCGTGGGCCATTGCTTTGTGGGCGGCTCCCAAGTT 2880
3090 GTTATCTCTGACGAGCTACGGCTGGGTGATCTGCTTCCCGCCCGCGGTATTGGGAG 3149

2881 GTTATCTCTGACGAGCTACGGCTGGGTGGATCTCTGCTTCCCGCGCGGTATTGGGAG 2940
3150 CTGCTGCTCAATACCGAGAGGTTCGCA CGCTGATCTCTTCCACCCACACCTTGGATGAG 3209
2941 CTGCTGCTCAATACCGAGAGGTTCGCA CGCTGATCTCTTCCACCCACACCTTGGATGAG 3000
3210 GCAGAGCTGCTGGAGACCGGTGGCTGTGGTGGCAGGTGGCGCTTGTCTGTCTGTGGC 3269
3001 GCAGAGCTGCTGGAGACCGGTGGCTGTGGTGGCAGGTGGCGCTTGTCTGTCTGTGGC 3060
3270 TCCCACTCTTCTTCCTGGCGCGTCA CCTGGGCTCCGGCTACTACCTGAGCGTGTGAAGGCC 3329
3061 TCCCACTCTTCTTCCTGGCGCGTCA CCTGGGCTCCGGCTACTACCTGAGCGTGTGAAGGCC 3120
3330 CGCTGCGCTCTGA CCAACCAATGAGAAGGTGACATCTGACATGAGAGGCGAGTGGACACC 3389
3121 CGCTGCGCTCTGA CCAACCAATGAGAAGGTGACATCTGACATGAGAGGCGAGTGGACACC 3180
3390 AGCAGGAAAAGAGAAATGSCAGCAGGGCAGAGAGTCGGCACTCTCTCAGCTGCTGGCC 3449
3181 AGCAGGAAAAGAGAAATGSCAGCAGGGCAGAGAGTCGGCACTCTCTCAGCTGCTGGCC 3240
3450 CTGGTACAGCACTGGGTGCCCCGGGCA CGGCTGGTGGAGGAGCTGCCACA CAGAGCTGGTG 3509
3241 CTGGTACAGCACTGGGTGCCCCGGGCA CGGCTGGTGGAGGAGCTGCCACA CAGAGCTGGTG 3300
3510 CTGGTCTGCCCTACACGGGTGCCATGACGGCAGCTTCGGCACA CACTTCCAGAGACTA 3569
3301 CTGGTCTGCCCTACACGGGTGCCATGACGGCAGCTTCGGCACA CACTTCCAGAGACTA 3360
3570 GACACGCGGTGCGGAGCTGAGGCTCAGCTGGCTACGGGATCTCCGACACAGCCTCGAG 3629
3361 GACACGCGGTGCGGAGCTGAGGCTCAGCTGGCTACGGGATCTCCGACACAGCCTCGAG 3420
3630 GAGATCTTCTGAAAGTGGTGGAGGAGTGTGCTGCGGACACA GATATGGAGGATGGCAGC 3689
3421 GAGATCTTCTGAAAGTGGTGGAGGAGTGTGCTGCGGACACA GATATGGAGGATGGCAGC 3480
3690 TGGGGGAGGACCTATGTCA CAGGCATTTCTGGGCTAGACGTAA CCGCTGGGCTCAAGATG 3749
3481 TGGGGGAGGACCTATGTCA CAGGCATTTCTGGGCTAGACGTAA CCGCTGGGCTCAAGATG 3540
3750 CCGCCACAGAGACAGCGCTGGAGAACGGGAA CACAGCTGGGTACGCCCCAGAGACTGAC 3809
3541 CCGCCACAGAGACAGCGCTGGAGAACGGGAA CACAGCTGGGTACGCCCCAGAGACTGAC 3600
3810 CAGGGCTCTGGGCGAGACGCGCTGGGCGGGGTA CAGGGCTGGGCACTGA CCGGCCAGCAG 3869
3601 CAGGGCTCTGGGCGAGACGCGCTGGGCGGGGTA CAGGGCTGGGCACTGA CCGGCCAGCAG 3660
3870 CTCAGGCGCTGCTTCTCAAGCGCTTTCGTGCTTGGCGCGAGCGCGCGGCGCTGTTTC 3929
3661 CTCAGGCGCTGCTTCTCAAGCGCTTTCGTGCTTGGCGCGAGCGCGCGGCGCTGTTTC 3720
3930 GCGCAGATGCTGCTGCTGCTTGTGGGCGCTGGCCCTCGTGTTCAGCCTCATCGTG 3989
3721 GCGCAGATGCTGCTGCTGCTTGTGGGCGCTGGCCCTCGTGTTCAGCCTCATCGTG 3780
3990 CTTCTTTTCGGGCACTACCGCGCTCTGCGGCTCAGTCCCA CCACTGTA CCGTGTCTCAGGTG 4049
3781 CTTCTTTTCGGGCACTACCGCGCTCTGCGGCTCAGTCCCA CCACTGTA CCGTGTCTCAGGTG 3840
4050 TCTTCTTTCAGTGAAGACGCCCCCAGGGGACCTTGGAGCGTGGCCCGGCTGCTCGAGGGCGCTG 4109
3841 TCTTCTTTCAGTGAAGACGCCCCCAGGGGACCTTGGAGCGTGGCCCGGCTGCTCGAGGGCGCTG 3900
4110 CTCAGGAGGACGACTGAGAGGAGCCCCCAGTGCAGCATAGCTCCCA CAGGTTCTCGGCA 4169
3901 CTCAGGAGGACGACTGAGAGGAGCCCCCAGTGCAGCATAGCTCCCA CAGGTTCTCGGCA 3960
4170 CCAGAGTTTCTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGGCAACTGGACCC CAGAGTCT 4229
3961 CCAGAGTTTCTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGGCAACTGGACCC CAGAGTCT 4020

6181	Db		CCGGAGGGGGCTGCGCCTTGGCGCGGTCTTTGGAGAGCTGGCGGTGCAACGGCGCAGAG	6240
6450	Qy		CAGGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGAGGAGGTAATCTTGTACTTC	6509
6241	Db		CAGGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGAGGAGGTAATCTTGTACTTC	6300
6510	Qy		TCCAGGACACAGGGGAAGGACGAGCACACGAGAGACAGAGGAGGACGAGGTGGGAGTG	6569
6301	Db		TCCAGGACACAGGGGAAGGACGAGCACACGAGAGACAGAGGAGGACGAGGTGGGAGTG	6360
6570	Qy		GACCCCGCGCAGGCCCTGACGACACCCCAAACGGGTACAGCCAGTTCTTCGATGACCCCTAGC	6629
6361	Db		GACCCCGCGCAGGCCCTGACGACACCCCAAACGGGTACAGCCAGTTCTTCGATGACCCCTAGC	6420
6630	Qy		ACTGCCGAGACTGTGCTGAGCCTCCTCCTCGCGGGCCGCGGGAGGCCCTGGGAA	6689
6421	Db		ACTGCCGAGACTGTGCTGAGCCTCCTCCTCGCGGGCCGCGGGAGGCCCTGGGAA	6480
6690	Qy		TGGCAAGGGCAAGGTAGAGTGCCCTAGAGCCCTGGACTCAGGCTGGCAGAGGGGCTGGTG	6749
6481	Db		TGGCAAGGGCAAGGTAGAGTGCCCTAGAGCCCTGGACTCAGGCTGGCAGAGGGGCTGGTG	6540
6750	Qy		CCCTGGAGAAAAATAAGAGAAGGCTGGAGAGAACCGTGCTGGTGAAA	6797
6541	Db		CCCTGGAGAAAAATAAGAGAAGGCTGGAGAGAACCGTGCTGGTGAAA	6588

RESULT 9

```

US-10-182-006-1
; Sequence 1, Application US/10182006
; Publication No. US20040048250A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: GENE ENCODING ABC-1 PARALOG AND THE POLYPEPTIDE DERIVED
; TITLE OF INVENTION: THEREFROM
; FILE REFERENCE: National Filing
; CURRENT APPLICATION NUMBER: US/10/182,006
; CURRENT FILING DATE: 2002-07-23
; PRIORITY APPLICATION NUMBER: PCT/US01/02191
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/177,889
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: 60/215,405
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6509)
; OTHER INFORMATION: Any
US-10-182-006-1

```

	Query Match	95.2%	Score 6478.4	DB 16	Length 6522
	Best Local Similarity	99.7%	Pred. No. 0		
	Matches 6501	Conservative 0	Mismatches 17	Indels 3	Gaps 1
Qy	210	ATGCGCTTCTGGACACAGCTGATGCTGCTGCTGTGGGAAGATTTCATGTATCGCCGGAGA	269		
Db	1	ATGCGCTTCTGGACACAGCTGATGCTGCTGTGGGAAGATTTCATGTATCGCCGGAGA	60		
Qy	270	CAGCGGTCACAGCTCCTGGTGCGAATTCGCTGTGGCCTCTCTTCCTCTTCTTCATCCTGGTG	329		
Db	61	CAGCGGTCACAGCTCCTGGTGCGAATTCGCTGTGGCCTCTCTTCCTCTTCTTCATCCTGGTG	120		
Qy	330	GCTGTTGCGCACTCCACACCGCCCTCGAGAGACCAATGAATGCCACTTCCCAAAACAGGCCA	389		
Db	121	GCTGTTGCGCACTCCACACCGCCCTCGAGAGACCAATGAATGCCACTTCCCAAAACAGGCCA	180		

Qy	390	CTGCCATCGCGGGCACCGTGCCTCGCTCCAGGGTCTCATCTGTATATGTGAACAACACC	449
Db	181	CTGCCATCGCGGGGCAACCGTGCCTCGCTCCAGGGTCTCATCTGTATATGTGAACAACACC	240
Qy	450	TGCTTTTCCGAGCTGACACCGGCGGAGGAGCCGCGGCGCTGTAGCAACTTCAACGACTCC	509
Db	241	TGCTTTTCCGAGCTGACACCGGCGGAGGAGCCGCGGCGCTGTAGCAACTTCAACGACTCC	300
Qy	510	CTGGTCTCCCGGCTGTAGCCGATGCGCACCTGTGTCTGGGAGGGGCCAGTGCCTCCACAGG	569
Db	301	CTGGTCTCCCGGCTGTAGCCGATGCGCGACCTGTGTCTGGGAGGGGCCAGTGCCTCCACAGG	360
Qy	570	ACGCTGCTGGCTTAGGGAAGCTGATGCGCACCGCTGAGGGCTGCAACGACGACGCGCCAG	629
Db	361	ACGCTGCTGGCTTAGGGAAGCTGATGCGCACCGCTGAGGGCTGCAACGACGACGCGCCAG	420
Qy	630	CCTCAACCAACCAAGCAGTCTTCCACTTGGAAACACCCATGCTGGATGTGCGGAGCTGCTG	689
Db	421	CCTCAACCAACCAAGCAGTCTTCCACTTGGAAACACCCATGCTGGATGTGCGGAGCTGCTG	480
Qy	690	ACGTCACTGCTGCGCAACCGAATCCCTGGGGTGTGGCACTGGGCCCAAGCCACGAGGACCTTGT	749
Db	481	ACGTCACTGCTGCGCAACCGAATCCCTGGGGTGTGGCACTGGGCCCAAGCCACGAGGACCTTGT	540
Qy	750	CACAGCTTGTGAGGCGCTGAGGACCTGGCCCGAGGAGCTCTGTGCGCTGCGCAGGCTG	809
Db	541	CACAGCTTGTGAGGCGCTGAGGACCTGGCCCGAGGAGCTCTGTGCGCTGCGCAGGCTG	600
Qy	810	GTGAGCTTCGGGCACTGCTGCAAGACCCGAGGAGACACGCGGCCCTCTGGAGTTGCTG	869
Db	601	GTGAGCTTCGGGCACTGCTGCAAGACCCGAGGAGACACGCGGCCCTCTGGAGTTGCTG	660
Qy	870	TCAGAGCCCTCTGACAGTGTACGGGACCTTAGCAGACAGTGTGGGCCCTCTCCCTCAAATGG	929
Db	661	TCAGAGCCCTCTGACAGTGTACGGGACCTTAGCAGACAGTGTGGGCCCTCTCCCTCAAATGG	720
Qy	930	TACAGGCTAGTGACCTGATGTGAGCTGTGTGGGCGAGGAGCCAGAAATCCGCGCTTCGCAGAC	989
Db	721	TACAGGCTAGTGACCTGATGTGAGCTGTGTGGGCGAGGAGCCAGAAATCCGCGCTTCGCAGAC	780
Qy	990	AGACGCTGAGCCCGGCTGCTCGGAGCTGATTTGGAGCCCTGACACAGCCACCCGCTGTCC	1049
Db	781	AGACGCTGAGCCCGGCTGCTCGGAGCTGATTTGGAGCCCTGACACAGCCACCCGCTGTCC	840
Qy	1050	CGCTGCTCTGAGAGCGCCTGAAGCCTCTGATCCTCGGGAAGCTACTCTTTTGACACCAAGAT	1109
Db	841	CGCTGCTCTGAGAGCGCCTGAAGCCTCTGATCCTCGGGAAGCTACTCTTTTGACACCAAGAT	900
Qy	1110	ACACCTTTTACCCGGAAGCTCATGGCCCAAGTGAAACCGACCTTTCGAGGAGCTCAACCTTG	1169
Db	901	ACACCTTTTACCCGGAAGCTCATGGCCCAAGTGAAACCGGACCTTTCGAGGAGCTCAACCTTG	960
Qy	1170	CTAGGAGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACTTCAATGAAC	1229
Db	961	CTAGGAGATGTCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACTTCAATGAAC	1020
Qy	1230	GACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCTCTGCAGATGACGAGTGAAGGAAGAAGG	1289
Db	1021	GACAGTTTCCAAATGTGGCCATGCTGCAGCGGCTCTCTGCAGATGACGAGTGAAGGAAGAAGG	1080
Qy	1290	CAGCCACAGACCTGGAGGCCGGGACCACTAGGAGGCCCTGCGATCTTTCTGGACCTCTGGG	1349
Db	1081	CAGCCACAGACCTGGAGGCCGGGACCACTAGGAGGCCCTGCGATCTTTCTGGACCTCTGGG	1140
Qy	1350	AGCGGTGGCTACAGCTGGGAGGACGACACGCTGTATGTGGGGCACTGTGTGGGCAACGCTG	1409
Db	1141	AGCGGTGGCTACAGCTGGGAGGACGACACGCTGTATGTGGGGCACTGTGTGGGCAACGCTG	1200
Qy	1410	GGCGGAGTGAAGAGTGCCTGTCTTGTGGAACAAGCTGGAGGCGGACCCCTCAGAGGAGGCC	1469
Db	1201	GGCGGAGTGAAGAGTGCCTGTCTTGTGGAACAAGCTGGAGGCGGACCCCTCAGAGGAGGCC	1260
Qy	1470	CTGTGTGTGCGCGGGCCCTGCAACTGTCTCGCGGAACATCGAATTTCTGGGCGGCGCTGCTTCT	1529

Db 1261 |||||CTGGTGTGCGGCGCTGCAACTGCTCGGGAACATCGAATCTGGGCGGGGCTGCTCTTC 1320
Qy 1530 |||||TTGGGACCTGAGACTCTTTAGACCCCAAGAGACCCCAACCCAGAGACTGGGCGCCCGG 1589
Db 1321 |||||TTGGGACCTGAGACTCTTTAGACCCCAAGAGACCCCAACCCAGAGACTGGGCGCCCGG 1380
Qy 1590 |||||CAGTGGGATCAAAATCCGATGACATTTAGAGTGGTCAAGAGACCAATAAGATCAGG 1649
Db 1381 |||||CAGTGGGATCAAAATCCGATGACATTTAGAGTGGTCAAGAGACCAATAAGATCAGG 1440
Qy 1650 |||||GACAGTTTGGGACCTTGGGCGGCGGAGCCCTGACCGAGCTGGCTACGTGTGG 1709
Db 1441 |||||GACAGTTTGGGACCTTGGGCGGCGGAGCCCTGACCGAGCTGGCTACGTGTGG 1500
Qy 1710 |||||GGCGCTTCTGTACTCTCAAGACCTGGTGGAGCGTGCAGCGCTCCGGTGTCTCAGCGGC 1769
Db 1501 |||||GGCGCTTCTGTACTCTCAAGACCTGGTGGAGCGTGCAGCGCTCCGGTGTCTCAGCGGC 1560
Qy 1770 |||||GCCAACCCCGGCGCGGCTCTTACCTGACGAGTGCCTATCTGCTGTATGTGAGCAGC 1829
Db 1561 |||||GCCAACCCCGGCGCGGCTCTTACCTGACGAGTGCCTATCTGCTGTATGTGAGCAGC 1620
Qy 1830 |||||GTGTTCTGCGGTGCTGAGCGGTGCTGCGGCTCTTCTGACGCTGGCGCTGGATCTAC 1889
Db 1621 |||||GTGTTCTGCGGTGCTGAGCGGTGCTGCGGCTCTTCTGACGCTGGCGCTGGATCTAC 1680
Qy 1890 |||||TCCGTGACATGACAGTGAAGCCCTGGTGGGAGAAAGAGACCGCGCTGGGAGACACC 1949
Db 1681 |||||TCCGTGACATGACAGTGAAGCCCTGGTGGGAGAAAGAGACCGCGCTGGGAGACACC 1740
Qy 1950 |||||ATGCGCGCATGGGCTCAGCGCGCGTGTCTGGCTAGGTGGTGTCTCAGCTGGCTC 2009
Db 1741 |||||ATGCGCGCGTGGGCTCAGCGCGCGTGTCTGGCTAGGTGGTGTCTCAGCTGGCTC 1800
Qy 2010 |||||GGGCGCTTCTGCTCAGCGCGCGCTGCTGGTCTTGGTCTCAAGCTGGGAGACATCTC 2069
Db 1801 |||||GGGCGCTTCTGCTCAGCGCGCGCTGCTGGTCTTGGTCTCAAGCTGGGAGACATCTC 1860
Qy 2070 |||||CCCTACAGCCACCGGGGCTGCTTCTGCTTCTGGGAGCCTTCCGGTGGCGCACGCTG 2129
Db 1861 |||||CCCTGACGACACCGGGGCTGCTTCTGCTTCTTGGCAGCTTCCGGTGGCGCACGCTG 1920
Qy 2130 |||||ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCTGGCGC 2189
Db 1921 |||||ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCCGCGCAACCTGGCTGGCGCACGCTG 1980
Qy 2190 |||||GGCTGGGCTACTTCTCCTCTACTGCGCTTACGTGCTGTGTGGCTTGGCGGAGCCGG 2249
Db 1981 |||||GGCTGGGCTACTTCTCCTCTACTGCGCTTACGTGCTGTGTGGCTTGGCGGAGCCGG 2040
Qy 2250 |||||CTGCGCGGGTGGCGGCTGGCGAGCCTGCTGCTGCGCGCTGGCGCTTGGCTTGGCG 2309
Db 2041 |||||CTGCGCGGGTGGCGGCTGGCGAGCCTGCTGCTGCGCGCTGGCGCTTGGCTTGGCG 2100
Qy 2310 |||||TGCGAGAGCTTGGCTCTGCTGAGAGCAGGCGAGGCGCGAGTGGCAACAGTGGCG 2369
Db 2101 |||||TGCGAGAGCTTGGCTCTGCTGAGAGCAGGCGAGGCGCGAGTGGCAACAGTGGCG 2160
Qy 2370 |||||ACCGGCGCTACGGCAGAGCTTCTAGCGCTGGCGCAGGTCTCTGGCTTCTGCTGGAC 2429
Db 2161 |||||ACCGGCGCTACGGCAGAGCTTCTAGCGCTGGCGCAGGTCTCTGGCTTCTGCTGGAC 2220
Qy 2430 |||||CGCGCGCTTACGGCTCGCCACCTGGTACCTGGAAGCTGTGCGCCAGGCGAGTACGGG 2489
Db 2221 |||||CGCGCGCTTACGGCTCGCCACCTGGTACCTGGAAGCTGTGCGCCAGGCGAGTACGGG 2280
Qy 2490 |||||ATCCCTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGGTGGGACCTCGGCGCCCG 2549
Db 2281 |||||ATCCCTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGGTGGGACCTCGGCGCCCG 2340
Qy 2550 |||||AAGAGTCTCAGCGCTTGGCCCAACCGCGCTGACCCCAAGGTGCTGGTAGAAGAGGCAACCG 2609

Db 2341 |||||AAGAGTCTCAGCGCTTGGCCCAACCGCGCTGGACCCAAAGGTGCTGGTAGAAGAGGACCG 2400
Qy 2610 |||||CCCGCTCTGAGTCTTGGGCTATCCGTTCCGAGCCTGGAGAAAGCTTTCTCTGAAAGCCCG 2669
Db 2401 |||||CCCGCTCTGAGTCTTGGGCTCTCGTTCCGAGCCTGGAGAAAGCTTTCTCTGAAAGCCCG 2460
Qy 2670 |||||CAGCAGCCCTGCGGGGCTCAGCTGAGCTTCTACAGGGCCAATCACCAGCTTCTCTG 2729
Db 2461 |||||CAGCAGCCCTGCGGGGCTCAGCTGAGCTTCTACAGGGCCAATCACCAGCTTCTCTG 2520
Qy 2730 |||||GGCCACAAAGGGCGGCAAGACCAACCTGCTCATCTTGGTGGCTCTTCTCCACCC 2789
Db 2521 |||||GGCCACAAAGGGCGGCAAGACCAACCTGCTCATCTTGGTGGCTCTTCTCCACCC 2580
Qy 2790 |||||AGTGGTGGCTCTGCGCTTCTATCTGGGCAACAGCTCCGCTCAGCATGGCGCCATCCGG 2849
Db 2581 |||||AGTGGTGGCTCTGCGCTTCTATCTGGGCAACAGCTCCGCTCAGCATGGCGCCATCCGG 2640
Qy 2850 |||||CCCAACCTGGGCTGTCTCTCAGTACAAAGTGTCTTGTGACATGCTGACCGTGGACGAG 2909
Db 2641 |||||CCCAACCTGGGCTGTCTCTCAGTACAAAGTGTCTTGTGACATGCTGACCGTGGACGAG 2700
Qy 2910 |||||CAGCTCTGTTCTATGGCGGCTGAAAGGTCTGAGTGGCGCTGAGTGGCGCCCGAGCAG 2969
Db 2701 |||||CAGCTCTGTTCTATGGCGGCTGAAAGGTCTGAGTGGCGCTGAGTGGCGCCCGAGCAG 2760
Qy 2970 |||||GACGCTCTGCTGACGAGTGTGGGCTGGTCTCCAAGCAGAGTGTGACAGCTCGCCACCTC 3029
Db 2761 |||||GACGCTCTGCTGACGAGTGTGGGCTGGTCTCCAAGCAGAGTGTGACAGCTCGCCACCTC 2820
Qy 3030 |||||TCTGGTGGATGCAACCGGAAGCTGTCTGCTGCCATTTGCTTTGTGGCGGCTCCCAAGTT 3089
Db 2821 |||||TCTGGTGGATGCAACCGGAAGCTGTCTGCTGCCATTTGCTTTGTGGCGGCTCCCAAGTT 2880
Qy 3090 |||||GTTATCTTGGACGAGCTTACGCGTGGCTGGTCTTCCCGCGCGGCTATTTGGAG 3149
Db 2881 |||||GTTATCTTGGACGAGCTTACGCGTGGCTGGTCTTCCCGCGCGGCTATTTGGAG 2940
Qy 3150 |||||CTGCTGCTCAATACCGAGAGGTGCGACGCTGATCTCTCCACCCACCTGATGAG 3209
Db 2941 |||||CTGCTGCTCAATACCGAGAGGTGCGACGCTGATCTCTCCAACCCACCTGATGAG 3000
Qy 3210 |||||GCAGAGCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGTGTGTGGC 3269
Db 3001 |||||GCAGAGCTGTGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTTGTGTGTGTGGC 3060
Qy 3270 |||||TCCCACTCTTCTGCGCGCTCACTGGGCTTCCGGCTACTACCTGACGCTGGTGAAGCC 3329
Db 3061 |||||TCCCACTCTTCTGCGCGCTCACTGGGCTTCCGGCTACTACCTGACGCTGGTGAAGCC 3120
Qy 3330 |||||CGCTGCGCTTACCAACAAATGAGAAGCTGACACTGACATGGAGGGCAGTGTGGACACC 3389
Db 3121 |||||CGCTGCGCTTACCAACAAATGAGAAGCTGACACTGACATGGAGGGCAGTGTGGACACC 3180
Qy 3390 |||||AGGCAGAAAAGAAATGGCAGCCAGGGCAGAGAGTGGCACTCTCTCAGCTGTGTGGCC 3449
Db 3181 |||||AGGCAGAAAAGAAATGGCAGCCAGGGCAGAGAGTGGCACTCTCTCAGCTGTGTGGCC 3240
Qy 3450 |||||CTGCTACAGCACTGGGTGCGCGGGCAGCGCTGGTGGAGGAGCTGCCACACGAGCTGTG 3509
Db 3241 |||||CTGCTACAGCACTGGGTGCGCGGGCAGCGCTGGTGGAGGAGCTGCCACACGAGCTGTG 3300
Qy 3510 |||||CTGCTGCTGCCCTACACCGGGTCCCATGACGCGAGCTTCCGCACTCTTCCGAGAGCTA 3569
Db 3301 |||||CTGCTGCTGCCCTACACCGGGTCCCATGACGCGAGCTTCCGCACTCTTCCGAGAGCTA 3360
Qy 3570 |||||GACAACCGGCTGGCGGAGCTGAGCTCACTGGCTACCGGATCTCCGACACAGGCTCGAG 3629
Db 3361 |||||GACAACCGGCTGGCGGAGCTGAGGCTCACTGGCTACCGGATCTCCGACACAGGCTCGAG 3420
Qy 3630 |||||GAGATCTTCTGAAGGTGGGAGAGTGTGCTCGGACACAGATATAGGAGATGGCAGC 3689
Db 3421 |||||GAGATCTTCTGAAGGTGGGAGAGTGTGCTCGGACACAGATATAGGAGATGGCAGC 3480

QY	3690	TGCGGCGACACCTATGTCAGAGCATTTGCTGGCTTAGACGTAAACCTTGGGCTCAAGATG	3749
Db	3481	TGCGGCGAGCACCTATGCAAGAGCATTTGCTGGCTTAGACGTAAACCTTGGGCTCAAGATG	3540
QY	3750	CCGCCACAGAGACAGCGCTGGAGAACCGGGAAACAGCTGGGTGAGCCGCCAGAGACTGAC	3809
Db	3541	CCGCCACAGAGACAGCGCTGGAGAACCGGGAAACAGCTGGGTGAGCCGCCAGAGACTGAC	3600
QY	3810	CAGGGCTCTGGGCGAGACGCCGTGGGCGGGTACAGGGCTGGGCACTGAACCCGCCAGCAG	3869
Db	3601	CAGGGCTCTGGGCGAGACGCCGTGGGCGGGTACAGGGCTGGGCACTGAACCCGCCAGCAG	3660
QY	3870	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCGAGCGCGCGGCGCTGTTTC	3929
Db	3661	CTCCAGGCCCTGCTTCTCAAGCGCTTCTGCTTGGCCCGCGAGCGCGCGGCGCTGTTTC	3720
QY	3930	GCCACAGATCGTGTGCTGCCCTCTTTGTGGGCGTGGCCCTCGTGTTCAGCCTCATCGTG	3989
Db	3721	GCCACAGATCGTGTGCTGCCCTCTTTGTGGGCGTGGCCCTCGTGTTCAGCCTCATCGTG	3780
QY	3990	CTCTCTTTCCGGCACTACCGGCTCTGCGGCTCAGTCCCAACCATGTAAGGTGCTCAGGTG	4049
Db	3781	CTCTCTTTCCGGCACTACCGGCTCTGCGGCTCAGTCCCAACCATGTAAGGTGCTCAGGTG	3840
QY	4050	TCCCTTCTTCACTAGGACGCCGCCAGGGACCTGGAGCTGCCCGGCTGCTCGAGGCGCTG	4109
Db	3841	TCCCTTCTTCACTAGGACGCCGCCAGGGACCTGGAGCTGCCCGGCTGCTCGAGGCGCTG	3900
QY	4110	CTGCAGGACGACGACTGGAGGAGCCCCCAGTCAGCATAGCTCCCAAGGTTCTCGGCA	4169
Db	3901	CTGCAGGACGACGACTGGAGGAGCCCCCAGTCAGCATAGCTCCCAAGGTTCTCGGCA	3960
QY	4170	CCAGAAAGTTCCCTGTAAGTGGCCAAAGTCTTGGCCAGTGGCAATGGACCCCGAGAGTCT	4229
Db	3961	CCAGAAAGTTCCCTGTAAGTGGCCAAAGTCTTGGCCAGTGGCAATGGACCCCGAGAGTCT	4020
QY	4230	CCATCCCCAGCTGCTAGCTAGCCAGCCCGGCTGCCCGGCGCTGCTGCCCGGCTGCCCG	4289
Db	4021	CCATCCCCAGCTGCTAGCTAGCCAGCCCGGCTGCCCGGCGCTGCTGCCCGGCTGCCCG	4080
QY	4290	GCTGACAGCTGGTGGTCCCCCTCCGCCCCAGGAGTGAACCGGCTCTGGGGAAAGTGGTTCAG	4349
Db	4081	GCTGACAGCTGGTGGTCCCCCTCCGCCCCAGGAGTGAACCGGCTCTGGGGAAAGTGGTTCAG	4140
QY	4350	AACTCTGACAGGCGGAACTGTCTGACTTCCCTGGTCAAGACCTACCGCGCCCTGGTGCGC	4409
Db	4141	AACTCTGACAGGCGGAACTGTCTGACTTCCCTGGTCAAGACCTACCGCGCCCTGGTGCGC	4200
QY	4410	CAGGGCTGAAGACTAAGAAAGTGGGTGAATGAGGTGAGGTACGGAGGCTTCTCGCTGGGG	4469
Db	4201	CAGGGCTGAAGACTAAGAAAGTGGGTGAATGAGGTGAGGTACGGAGGCTTCTCGCTGGGG	4260
QY	4470	GGCGGAGACCCAGGCTGCGCTCGGGCCAGAGTTGGGCGGCTCAGTGGAGGAGTGTGG	4529
Db	4261	GGCGGAGACCCAGGCTGCGCTCGGGCCAGAGTTGGGCGGCTCAGTGGAGGAGTGTGG	4320
QY	4530	GCCTGCTGAGTCCCCTGCTGCGGGGCCCTCGA CCGTGTCTGTAAGAAACCTCACAGCC	4589
Db	4321	GCCTGCTGAGTCCCCTGCTGCGGGGCCCTCGA CCGTGTCTGTAAGAAACCTCACAGCC	4380
QY	4590	TGGGCTCAAGCCTGGATGCTCAGGACAGTCTCAAGTCTGTTTCAACAAGGCTGG	4649
Db	4381	TGGGCTCAAGCCTGGATGCTCAGGACAGTCTCAAGTCTGTTTCAACAAGGCTGG	4440
QY	4650	CAGTCAATGGTGGCTTTGTCAACCCAGCCAGCAACGCAATCCCTGGTGTCACTGGCCC	4709
Db	4441	CAGTCAATGGTGGCTTTGTCAACCCAGCCAGCAACGCAATCCCTGGTGTCACTGGCCC	4500
QY	4710	CCAGGCGCGGCGCCGACAGCCCAAGCATCAACCACTCAACCCCTTTGAAACCTCAACC	4769
Db	4501	CCAGGCGCGGCGCCGACAGCCCAAGCATCAACCACTCAACCCCTTTGAAACCTCAACC	4560

QY	4770	AAGGACAGCTGTCTGAGGCTGCACTGATGGGCTCCTCGGTGGACGCTCCTGCTCTCCATC	4829
Db	4561	AAGGACAGCTGTCTGAGGCTGCACTGATGGGCTCCTCGGTGGACGCTCCTGCTCTCCATC	4620
QY	4830	TGTGTGGTCTTTGGCCATGTCTTTGTCCCGGCGAGTTTCACTTGTCTCTCATTTGAGGAG	4889
Db	4621	TGTGTGGTCTTTGGCCATGTCTTTGTCCCGGCGAGTTTCACTTGTCTCTCATTTGAGGAG	4680
QY	4890	CGAGTCAACCGGACGCAAGCACCTGCGAGCTCATGGGGGGGCTGTCCCGCACCTCTACTGG	4949
Db	4681	CGAGTCAACCGGACGCAAGCACCTGCGAGCTCATGGGGGGGCTGTCCCGCACCTCTACTGG	4740
QY	4950	CTTTGGCAACTTTCTTGGGACATGTGTAACTACTTGTGTGCGACATGCACTGTGTGTGCTC	5009
Db	4741	CTTTGGCAACTTTCTTGGGACATGTGTAACTACTTGTGTGCGACATGCACTGTGTGTGCTC	4800
QY	5010	ATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGTCACCACTGTCTGTCTCTCTG	5069
Db	4801	ATCTTTCTGGCCCTTCCAGCAGAGGGCATATGTGGCCCTGTCACCACTGTCTGTCTCTCTG	4860
QY	5070	CTGTGCTACTACTGTATGCGTGGTGCATCACACGCTCATGTATCCAGGCTCTCTCTTTC	5129
Db	4861	CTGTGCTACTACTGTATGCGTGGTGCATCACACGCTCATGTATCCAGGCTCTCTCTTTC	4920
QY	5130	TTCTCCGTGCCAGCAGACGCTATGTGTGTGCTCACCTGTCATAAACTCTTTATTGGCATC	5189
Db	4921	TTCTCCGTGCCAGCAGACGCTATGTGTGTGCTCACCTGTCATAAACTCTTTATTGGCATC	4980
QY	5190	AATGGAAGCATGSCCACTTTGTGTGAGCTCTTCTCTGATCAGAACTGTGAGGAGGTG	5249
Db	4981	AATGGAAGCATGSCCACTTTGTGTGAGCTCTTCTCTGATCAGAACTGTGAGGAGGTG	5040
QY	5250	AGCGGATCTTGAACACAGCTCTTCTTATCTTCCCGCATCTTCTGCTTGGGCGGGGGCTC	5309
Db	5041	AGCGGATCTTGAACACAGCTCTTCTTATCTTCCCGCATCTTCTGCTTGGGCGGGGGCTC	5100
QY	5310	ATTGACATGCTGGGAAACGAGGCAATGGCTGTGCTTTGAGCGCTTGGAGACAGGCGAG	5369
Db	5101	ATTGACATGCTGGGAAACGAGGCAATGGCTGTGCTTTGAGCGCTTGGAGACAGGCGAG	5160
QY	5370	TTCTCAGTCACTCCCTGCGCTGGGAGGTGGTGGCAAGAACTCTTTGCGCAATGTAACAG	5429
Db	5161	TTCTCAGTCACTCCCTGCGCTGGGAGGTGGTGGCAAGAACTCTTTGCGCAATGTAACAG	5220
QY	5430	GGGCCCCCTTCTCTTCTTCTTCACTACTGCTGTCAGCA CCGAGGCAAACTCTCTGCCACAG	5489
Db	5221	GGGCCCCCTTCTCTTCTTCTTCACTACTGCTGTCAGCA CCGAGGCAAACTCTCTGCCACAG	5280
QY	5490	CCGAGGCTGAGGTCTCTGCACTCTCTGGGAGGAGGAGGATGAGCCCGTGAACGG	5549
Db	5281	CCGAGGCTGAGGTCTCTGCACTCTCTGGGAGGAGGAGGATGAGCCCGTGAACGG	5340
QY	5550	GAGCGGCTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTGTCTGAGGAACTTTGACCAAG	5609
Db	5341	GAGCGGCTGGTCCAAAGGAGCCACCCAGGGGGATGTGTGGTGTCTGAGGAACTTTGACCAAG	5400
QY	5610	GTAATACGCTGGGAGAGGATGCGCAGCTGTGACCGCTTGTGCTGCTGGGATTTCCCTTGGT	5669
Db	5401	GTAATACGCTGGGAGAGGATGCGCAGCTGTGACCGCTTGTGCTGCTGGGATTTCCCTTGGT	5460
QY	5670	GAGTGTGTTGGGCTGTGGGTGTGAATGGAGGAGGAAACGCTCCACGTTTCCGATGGTG	5729
Db	5461	GAGTGTGTTGGGCTGTGGGTGTGAATGGAGGAGGAAACGCTCCACGTTTCCGATGGTG	5520
QY	5730	ACGGGGGACACATTTGGCCAGCAGGGGCGAGGCTGTCTGTGGCAGGCGCACAGCTGGGCCGG	5789
Db	5521	ACGGGGGACACATTTGGCCAGCAGGGGCGAGGCTGTCTGTGGCAGGCGCACAGCTGGGCCGG	5580
QY	5790	GAAACCCAGTGTGCGGACCTCTACATGCGGATGCTGCGCTCAATCCGATGCCATCTTTGAG	5849
Db	5581	GAAACCCAGTGTGCGGACCTCTACATGCGGATGCTGCGCTCAATCCGATGCCATCTTTGAG	5640
QY	5850	CTGCTGACGGGCGCGAGCACCTGGAGCTGCTTGGCGGCTGCGCGGTGTCTCCGGAGGCC	5909

```
Db 5641 CTGCTGACGGGCGGAGCACCTGAGCTGCTTGGCGGCTCGCGGCTGTCGGAGGCC 5700
Qy 5910 CAGGTTGCCAGACCGCTGGCTCAGCGCTGGCGGCTCTGGGACTCTCATGGTACGACAG 5969
Db 5701 CAGGTTGCCAGACCGCTGGCTCAGCGCTGGCGGCTCTGGGACTCTCATGGTACGACAG 5760
Qy 5970 CGGCTTCAGGACCTACAGCGAGAGGAAACAAACGAAAGTGGCGACCGCCCTGGCGCTG 6029
Db 5761 CGGCTTCAGGACCTACAGCGAGAGGAAACAAACGAAAGTGGCGAGCGCCCTGGCGCTG 5820
Qy 6030 GTTGGGACCCAGCGCTGGTGTCTTGACAGCGCCGACCAAGGATGACACCCAGCGG 6089
Db 5821 GTTGGGACCCAGCGCTGGTGTCTTGACAGCGCCGACCAAGGATGACACCCAGCGG 5880
Qy 6090 CGGCGCTTCTTTGGAACAGACCTTTTGGCGGTGGTGGGAGGCGCTTCAGTGAATGTC 6149
Db 5881 CGGCGCTTCTTTGGAACAGACCTTTTGGCGGTGGTGGGAGGCGCTTCAGTGAATGTC 5940
Qy 6150 ACCTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGAAT 6209
Db 5941 ACCTCCCATAGCATGAGGAGTGTGAAGCGCTCTGCTCGCGCTAGCCATCATGTGAAT 6000
Qy 6210 GGGCGGTTCCGCTGCTGGGACGCGCCGCAACATCTCAAGGGCAGATTTCGGCGGGTCA 6269
Db 6001 GGGCGGTTCCGCTGCTGGGACGCGCCGCAACATCTCAAGGGCAGATTTCGGCGGGTCA 6060
Qy 6270 ACACGTGACCTTCGCGGTGCGCGGAGGTCGCGGAGGCAATGAGGCGCGCTTCGTTGCGG 6329
Db 6061 ACACGTGACCTTCGCGGTGCGCGGAGGTCGCGGAGGCAATGAGGCGCGCTTCGTTGCGG 6120
Qy 6330 GAGTTCCCTGGTTCGAGGAGTGTGAGGAGGCAATGAGGCGCGCTTCGTTGCGGAGG 6389
Db 6121 GAGTTCCCTGGTTCGAGGAGTGTGAGGAGGCAATGAGGCGCGCTTCGTTGCGGAGG 6180
Qy 6390 CGGAGGCGGCTGCGGCTGCGCGGAGGTCGCGGAGGCAATGAGGCGCGCTTCGTTGCGG 6449
Db 6181 CGGAGGCGGCTGCGGCTGCGCGGAGGTCGCGGAGGCAATGAGGCGCGCTTCGTTGCGG 6240
Qy 6450 CAGCGCTGGAGGACTTTTCCGTGAGCCAGACGATCTTGGAGAGGAGGATTTCTGTACTTC 6509
Db 6241 CAGCGCTGGAGGACTTTTCCGTGAGCCAGACGATCTTGGAGAGGAGGATTTCTGTACTTC 6300
Qy 6510 TCCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6569
Db 6301 TCCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6360
Qy 6570 GACCCCGCGCAGGCTGCGAGGACCCCAACCGGTCAGCCAGTTCCTCGATGACCCCTAGC 6629
Db 6361 GACCCCGCGCAGGCTGCGAGGACCCCAACCGGTCAGCCAGTTCCTCGATGACCCCTAGC 6420
Qy 6630 ACTCGGAGACTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTG 6686
Db 6421 ACTCGGAGACTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTG 6480
Qy 6687 GAATGGCAAGGCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 6727
Db 6481 GAATGGCAAGGCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 6521
```

RESULT 10

```
US-09-858-194-3
; Sequence 3, Application US/09858194
; Patent No. US2002061590A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, MARIA
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: 38594, A NOVEL HUMAN TRANSPORTER AND USES THEREOF
; FILE REFERENCE: MNI-153
; CURRENT APPLICATION NUMBER: US/09/858,194
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/204,211
; PRIOR FILING DATE: 2000-05-12
```

```
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6432)
US-09-858-194-3
```

```
Query Match 94.0%; Score 6398.4; DB 9; Length 6432;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6421; Conservative 0; Mismatches 11; Indels 6; Gaps 1;
```

```
Qy 210 ATGGCCCTTCTCGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCCGAGA 269
Db 1 ATGGCCCTTCTCGACACAGCTGATGCTGCTCTGGAAGAAATTTTCATGTATCGCCGAGA 60
Qy 270 CAGCGGTCACAGCTCCTGGTCGAATTGCTGTGGCCCTCTCTTCCTCTTTCATCTCTGGTG 329
Db 61 CAGCC-----GCTCCTGGTCGAATTGCTGTGGCCCTCTCTTCCTCTTTCATCTCTGGTG 114
Qy 330 GCTGTTCGCCACTCCACCCGCCCTTGAGCAACATGAATGCCATTTCCCAAAACAAGCCA 389
Db 115 GCTGTTCGCCACTCCACCCGCCCTTGAGCAACATGAATGCCATTTCCCAAAACAAGCCA 174
Qy 390 CTGCATCGGCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449
Db 175 CTGCATCGGCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 234
Qy 450 TGCTTTCCGAGCTGACACCGGGGAGGAGCGCGGCGCTGAGCAACTTTCAACGACTCC 509
Db 235 TGCTTTCCGAGCTGACACCGGGGAGGAGCGCGGCGCTGAGCAACTTTCAACGACTCC 294
Qy 510 CTGTCTCCCGCTGCTAGCCGATGCCGCACTGTGTGGAGGGGCGAGTGCCTCAAGG 569
Db 295 CTGTCTCCCGCTGCTAGCCGATGCCGCACTGTGTGGAGGGGCGAGTGCCTCAAGG 354
Qy 570 ACGTGTGCTGCTAGGAGAGCTGATGCCAGCTGAGGGCTGACGAGCAACGCGCCAG 629
Db 355 ACGTGTGCTGCTAGGAGAGCTGATGCCAGCTGAGGGCTGACGAGCAACGCGCCAG 414
Qy 630 CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCACTGTGTGGAGTGTGCGGAGCTGTG 689
Db 415 CCTCAACCAACCAAGCAGTCTCCACTGGAACCAACCACTGTGTGGAGTGTGCGGAGCTGTG 474
Qy 690 ACGTCACTGTGCGCAACGAAATCCCTGGGGTTGGCACTGGGCCCAAGCCAGAGGCCCTTG 749
Db 475 ACGTCACTGTGCGCAACGAAATCCCTGGGGTTGGCACTGGGCCCAAGCCAGAGGCCCTTG 534
Qy 750 CACAGCTTGTGGAGGCGCTGAGGACCTGGCCAGGAGCTCTGGCGCTGGCAGGCTG 809
Db 535 CACAGCTTGTGGAGGCGCTGAGGACCTGGCCAGGAGCTCTGGCGCTGGCAGGCTG 594
Qy 810 GTGAGCTTGGGCACTGTGTCAGAGACCCCGAGGAGCCAGCGGCGCTGGAGTTGCTG 869
Db 595 GTGAGCTTGGGCACTGTGTCAGAGACCCCGAGGAGCCAGCGGCGCTGGAGTTGCTG 654
Qy 870 TCAGAGGCCCTCTGCACTGTGTCAGGAGCCTAGCAGCAGTGGGGCCCTCCCTCAACTGG 929
Db 655 TCAGAGGCCCTCTGCACTGTGTCAGGAGCCTAGCAGCAGTGGGGCCCTCCCTCAACTGG 714
Qy 930 TACGAGGCTAGTGACCTGATGGAGCTGGTGGGAGGAGCCAGAAATCGCCCTGCCAGAC 989
Db 715 TACGAGGCTAGTGACCTGATGGAGCTGGTGGGAGGAGCCAGAAATCGCCCTGCCAGAC 774
Qy 990 AGCAGCTGAGCCCGCTGCTGCTGAGCTGATTCGAGCCCTGGAGCAGCAGCCGCTGCC 1049
Db 775 AGCAGCTGAGCCCGCTGCTGCTGAGCTGATTCGAGCCCTGGAGCAGCAGCCGCTGCC 834
Qy 1050 CGCCTGCTGCGAGACCGCTGAGCCCTGATCTCTCGGAGAGCTACTCTTTCACACAGAT 1109
```

Db 835 CGCCTGCTCTGGAGACGCGCTTGAAGCCTCTGATCCTCGGGAAGCTACTCTTTTGACACAGAT 894
Qy 1110 ACACCTTTTACCCGGAAGCTCTATGGCCCAAGTGAA CCGGACCTTCGAGGAGCTCACCCCTG 1169
Db 895 ACACCTTTTACCCGGAAGCTCATGGCCCAAGTGAA CCGGACCTTCGAGGAGCTCACCCCTG 954
Qy 1170 CTGAGGAGATGTCGGGAGGTGGGAGATGCTGGGACCCCGGATCTTTCACCTTCATGAAC 1229
Db 955 CTGAGGAGATGTCGGGAGGTGGGAGATGCTGGGACCCCGGATCTTTCACCTTCATGAAC 1014
Qy 1230 GACAGTTTCCAATGTGGGCATGCTGACGGCTCTCTGAGATGACAGATGAAGGAAGG 1289
Db 1015 GACAGTTTCCAATGTGGGCATGCTGACGGCTCTCTGAGATGACAGATGAAGGAAGG 1074
Qy 1290 CAGCCAGACCTGAGGCGCGGAGACCATGAGGACCTGCGATCTCTTTCTGGAACCTTGGG 1349
Db 1075 CAGCCAGACCTGAGGCGCGGAGACCATGAGGACCTGCGATCTCTTTCTGGAACCTTGGG 1134
Qy 1350 AGCGGTGGCTACAGCTGGCAGGAGCGACACGCTGATGCTGGGACCTGGTGGGACGCTG 1409
Db 1135 AGCGGTGGCTACAGCTGGCAGGAGCGACACGCTGATGCTGGGACCTGGTGGGACGCTG 1194
Qy 1410 GCGGAGTGACGAGTGCTCTGCTTGGACAAGCTGAGGCGGACCCCTCAGAGGACGCC 1469
Db 1195 GCGGAGTGACGAGTGCTCTGCTTGGACAAGCTGAGGCGGACCCCTCAGAGGACGCC 1254
Qy 1470 CTGGTGTGCGGGCCCTGGAACCTGCTCGCGGAACATGCGATCTGGGGCGGCGTGTCTTC 1529
Db 1255 CTGGTGTGCGGGCCCTGGAACCTGCTCGCGGAACATGCGATCTGGGGCGGCGTGTCTTC 1314
Qy 1530 TTGGGACCTGAGGACTCTTCAGACCCACAGACACCAACCCAGACCTGGGCCCCGGC 1589
Db 1315 TTGGGACCTGAGGACTCTTCAGACCCACAGACACCAACCCAGACCTGGGCCCCGGC 1374
Qy 1590 CAGCTGCGCATCAAAATCCGCATGGAATTCAGCTGTGTCAGAGGACCAATAAGATCAGG 1649
Db 1375 CAGCTGCGCATCAAAATCCGCATGGAATTCAGCTGTGTCAGAGGACCAATAAGATCAGG 1434
Qy 1650 GACAGGTTTGGGACCCCTGGCCAGCCGCGGAACCCCTGACCGACCTGGCGCTACGTGTGG 1709
Db 1435 GACAGGTTTGGGACCCCTGGCCAGCCGCGGAACCCCTGACCGACCTGGCGCTACGTGTGG 1494
Qy 1710 GCGGCTTCCTGATACCTGCAAGACTGCTGGAGCGTGACCGCTGCGCGCTCAGCGGC 1769
Db 1495 GCGGCTTCCTGATACCTGCAAGACTGCTGGAGCGTGACCGCTGCGCGCTCAGCGGC 1554
Qy 1770 GCCAAACCCCGGGCGCGCTCTACCTGCAGCAGATGCCCTATCCGCTGCTATGTGGACGAC 1829
Db 1555 GCCAAACCCCGGGCGCGCTCTACCTGCAGCAGATGCCCTATCCGCTGCTATGTGGACGAC 1614
Qy 1830 GTGTTCTGCTGTGTGAGCGGCTGCTGCGCTCTTCTGAGCTGGCTGATCTAC 1889
Db 1615 GTGTTCTGCTGTGTGAGCGGCTGCTGCGCTCTTCTGAGCTGGCTGATCTAC 1674
Qy 1890 TCCGTGACACTGACAGTGAAAGCGGCTGGGAGAGGAGAGCGCGCTGCGGACACC 1949
Db 1675 TCCGTGACACTGACAGTGAAAGCGGCTGGGAGAGGAGAGCGCGCTGCGGACACC 1734
Qy 1950 ATGCGGCCCATGGGGCTCAGCGCGCGGCTGCTCTGGCTAGGCTGGTTCTCAGCTGCGCTC 2009
Db 1735 ATGCGGCCCATGGGGCTCAGCGCGCGGCTGCTCTGGCTAGGCTGGTTCTCAGCTGCGCTC 1794
Qy 2010 GGGCCCTTCCTGCTCAGCGCGCGGCTGCTGCTGTGCTGATCTGAAGCTGGGGACATCCTC 2069
Db 1795 GGGCCCTTCCTGCTCAGCGCGCGGCTGCTGCTGTGCTGATCTGAAGCTGGGGACATCCTC 1854
Qy 2070 CCTTACAGCCACCCCGGCGGTGCTTCTGCTTCTTGGACCTTCGCGGTGGCCACGGTG 2129
Db 1855 CCTTACAGCCACCCCGGCGGTGCTTCTGCTTCTTGGACCTTCGCGGTGGCCACGGTG 1914
Qy 2130 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCGCGCCCAACCTGGCTGGCGGCTGGCGC 2189
Db 1915 ACCCAGAGCTTCTGCTCAGCGCTTCTTCTCCGCGCCCAACCTGGCTGGCGGCTGGCGC 1974

Qy 2190 GGCCTGGCCTACTTCTCCCTCTACCTGCCCCTACGTGCTGTGTGCTGGCTTGGCGGGACCGG 2249
Db 1975 GGCCTGGCCTACTTCTCCCTCTACCTGCCCCTACGTGCTGTGTGTGGCTTGGCGGGACCGG 2034
Qy 2250 CTGCCCGCGGGTGGCGCGCTGGCCGCGAGCTGCTGTGCGCCCTGGGCTTTCGGCTTCGGC 2309
Db 2035 CTGCCCGCGGGTGGCGCGCTGGCCGCGAGCTGCTGTGCGCCCTGGGCTTTCGGCTTCGGC 2094
Qy 2310 TGCAGAGCCTGCTCTGTGTGAGGAGCAGGGCGAGGGCGCGCAGGTGGGCAACAAGTGGGC 2369
Db 2095 TGCAGAGCCTGCTCTGTGTGAGGAGCAGGGCGAGGGCGCGCAGGTGGGCAACAAGTGGGC 2154
Qy 2370 ACCGGCTTACGCGACAGCTTTCAGCTGGCCCGCAGGTCTCTGCGCTTCTGCTGCTGGAC 2429
Db 2155 ACCGGCTTACGCGACAGCTTTCAGCTGGCCCGCAGGTCTCTGCGCTTCTGCTGCTGGAC 2214
Qy 2430 GCGGCGCTCTACGCGCTCGCCACCTGCTACCTGGAAGCTGTGTGCCCAGGCCAGTACGGG 2489
Db 2215 GCGGCGCTCTACGCGCTCGCCACCTGCTACCTGGAAGCTGTGTGCCCAGGCCAGTACGGG 2274
Qy 2490 ATCCCTGAACCATGGAATTTTCTTTTCGAGAGAGCTA CTGGTGGGACCTTCGGCCCCCC 2549
Db 2275 ATCCCTGAACCATGGAATTTTCTTTTCGAGAGAGCTA CTGGTGGGACCTTCGGCCCCCC 2334
Qy 2550 AAGAGTCCAGCCCTTTGCCCCACCCCGCTGGACCCCAAAGTGTGTGTAAGAAGGACCG 2609
Db 2335 AAGAGTCCAGCCCTTTGCCCCACCCCGCTGGACCCCAAAGTGTGTGTAAGAAGGACCG 2394
Qy 2610 CCGGCGCTCAGTCTCGCGCTATCCGTTTCGAGCCTGGAGAGCGCTTTTCCTGGAAGCCCG 2669
Db 2395 CCGGCGCTCAGTCTCGCGCTATCCGTTTCGAGCCTGGAGAGCGCTTTTCCTGGAAGCCCG 2454
Qy 2670 CAGCAGCCCTGCGGGGCTCAGCCTGGAATTTCTTACAGGGCCACATCA CCGCCTTCCTG 2729
Db 2455 CAGCAGCCCTGCGGGGCTCAGCCTGGAATTTCTTACAGGGCCACATCA CCGCCTTCCTG 2514
Qy 2730 GGCACAA CCGGGCCGCGGACCAACCCCTGCTCACTCTGAGTGGCCCTTTCGCCACCC 2789
Db 2515 GGCACAGCGGGGCCGCGCAAGACCAACCCCTGCTCACTCTGAGTGGCCCTTTCGCCACCC 2574
Qy 2790 AGTGTGGCTCTGCCCTTCACTCTCGGCGCA CAGCCTCGCTCCAGCATGGCGGCATCCGG 2849
Db 2575 AGTGTGGCTCTGCCCTTCACTCTCGGCGCA CAGCCTCGCTCCAGCATGGCGGCATCCGG 2634
Qy 2850 CCCACCTTGGGCTGTGTCTCAGTACAA CAGTGTGTTTGA CATGTGA CCGTGGACGAG 2909
Db 2635 CCCACCTTGGGCTGTGTCTCAGTACAA CAGTGTGTTTGA CATGTGA CCGTGGACGAG 2694
Qy 2910 CAGTCTGTTCTATGCGGCGCTGAAGGCTCTGAGTGCCTGTGTGGGCCCCGAGCAG 2969
Db 2695 CAGTCTGTTCTATGCGGCGCTGAAGGCTCTGAGTGCCTGTGTGGGCCCCGAGCAG 2754
Qy 2970 GACGCTGTGTCAGGATGTGGGCTGTGTC CAAAGCAGAGTGTGAGATCTCGGCACCTC 3029
Db 2755 GACGCTGTGTCAGGATGTGGGCTGTGTC CAAAGCAGAGTGTGAGATCTCGGCACCTC 2814
Qy 3030 TCTGTGGGATGCAA CCGAAGCTGTCCGTGG CCAATTGCTTTTGTGGGCGGCTCCAAAGTT 3089
Db 2815 TCTGTGGGATGCAA CCGAAGCTGTCCGTGG CCAATTGCTTTTGTGGGCGGCTCCCAAGTT 2874
Qy 3090 GTTATCTTGAACGAGCTTACGGCTGGGTGGATCTGCTTCCCGCGCGCTATTTGGGAG 3149
Db 2875 GTTATCTTGAACGAGCTTACGGCTGGGTGGATCTGCTTCCCGCGCGCTATTTGGGAG 2934
Qy 3150 CTGCTGTCAAATAC CAGAGAGGTTCGCA CCGCTGATCTCTCCACCCACACCTGGATGAG 3209
Db 2935 CTGCTGTCAAATAC CAGAGAGGTTCGCA CCGCTGATCTCTCCACCCACACCTGGATGAG 2994
Qy 3210 GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGAGGTGGCGGCTTGTGCTGTGGC 3269
Db 2995 GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGAGGTGGCGGCTTGTGCTGTGGC 3054

QY	3270	TCGCCACATCTTCTCGCGCGCTCACTTGGGCTCCGGCTACTACCTGAA	CGCTGGTGAAGGCC	3329
DB	3055	TCGCCACATCTTCTCGCGCGCTCACTTGGGCTCCGGCTACTACCTGAA	CGCTGGTGAAGGCC	3114
QY	3330	CGCCTGCCCTTGACCA	CCAAATGAGAAAGGCTGACACTGACATGGAGGCGAGTGTGGACACC	3389
DB	3115	CGCCTGCCCTTGACCA	CCAAATGAGAAAGGCTGACACTGACATGGAGGCGAGTGTGGACACC	3174
QY	3390	AGGCAGGAAAAAGAAATGGCAGCCAGGGCAGCAGAGTTCGGCACT	TCCTCAGCTGCTGGCC	3449
DB	3175	AGGCAGGAAAAAGAAATGGCAGCCAGGGCAGCAGAGTTCGGCACT	TCCTCAGCTGCTGGCC	3234
QY	3450	CTGGTTACAGCACTGGGTGCCCGGGGCACGGCTGCTGAGAGAGCTG	CCACACGAGCAGCTGGT	3509
DB	3235	CTGGTTACAGCACTGGGTGCCCGGGGCACGGCTGCTGAGAGAGCTG	CCACACGAGCAGCTGGT	3294
QY	3510	CTGGTGTCGCCCTACACGGGTGCCATGACGCGCAGCTTCGCCCA	CACATCTTCCGAGAGCTA	3569
DB	3295	CTGGTGTCGCCCTACACGGGTGCCATGACGCGCAGCTTCGCCCA	CACATCTTCCGAGAGCTA	3354
QY	3570	GACACGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCC	GACACACGAGCCTCGAG	3629
DB	3355	GACACGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCC	GACACACGAGCCTCGAG	3414
QY	3630	GAGATCTTCTGAAGGTGGTGGAGGAGTGTCTCGGACACAGATATG	GAAGGATGGGACG	3689
DB	3415	GAGATCTTCTTGAAGGTGGTGGAGGAGTGTCTCGGACACAGATATG	GAAGGATGGGACG	3474
QY	3690	TGCGGGCAGCACCTATGCACAGGCAATTCCTGGCCCTAGACGTMA	CCCTGCGGCTCAGATG	3749
DB	3475	TGCGGGCAGCACCTATGCACAGGCAATTCCTGGCCCTAGACGTMA	CCCTGCGGCTCAGATG	3534
QY	3750	CCGCCACAGGAGACAGCGCTGGGAAACGGGGAAACACAGCTGGG	TGAGCTCAGCCCCCAGAGACTGAC	3809
DB	3535	CCGCCACAGGAGACAGCGCTGGGAAACGGGGAAACACAGCTGGG	TGAGCTCAGCCCCCAGAGACTGAC	3594
QY	3810	CAGGGCTCTGGGCGACAGCGCTGGGCGCGGTATCAGGGCTGGGGA	CTGACCCCGCCAGCAG	3869
DB	3595	CAGGGCTCTGGGCGACAGCGCTGGGCGCGGTATCAGGGCTGGGGA	CTGACCCCGCCAGCAG	3654
QY	3870	CTCCAGGCCCCTTCTCAGCGCTTTCTGCTTGGCCCGCGCAGCGCG	CGCGGCGCTGTTTC	3929
DB	3655	CTCCAGGCCCCTTCTCAGCGCTTTCTGCTTGGCCCGCGCAGCGCG	CGCGGCGCTGTTTC	3714
QY	3930	GCCCAGATCTGTGCTGCCCTCTTTGTGGGCTTGCCCTCGTGTGTC	AGCCCTCATCTGTCG	3989
DB	3715	GCCCAGATCTGTGCTGCCCTCTTTGTGGGCTTGCCCTCGTGTGTC	AGCCCTCATCTGTCG	3774
QY	3990	CCTCCTTTGGGCACTACCCGGCTCTCGGGCTCAGTCCCAACCATAT	ACGTGTCTCAGGTG	4049
DB	3775	CCTCCTTTGGGCACTACCCGGCTCTCGGGCTCAGTCCCAACCATAT	ACGTGTCTCAGGTG	3834
QY	4050	TCCTTCTTCAGTGAAGAGCGCCACAGGGGACCTTGAGCGTCCGGCT	GTCTCGAGGCGCTG	4109
DB	3835	TCCCTTCTTCAGTGAAGAGCGCCACAGGGGACCTTGAGCGTCCGG	CTGTCTCGAGGCGCTG	3894
QY	4110	CTGCAGGAGGACAGCACTTGGAGGACGCCCACTAGTCAGCATAGCT	CCACAGGTTCTCGGCA	4169
DB	3895	CTGCAGGAGGACAGCACTTGGAGGAGCCCCCACTAGTCAGCATAGCT	CCACAGGTTCTCGGCA	3954
QY	4170	CCAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACT	GGACCCCGCAGTCT	4229
DB	3955	CCAGAGTTCCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAACT	GGACCCCGCAGTCT	4014
QY	4230	CCATCCCGAGCCTGCCAGTGAAGCAGCCCGGTGCTCCGGCGCTGTCT	GTGCCGACTGCCCG	4289
DB	4015	CCATCCCGAGCCTGCCAGTGAAGCAGCCCGGTGCTCCGGCGCTGTCT	GTGCCGACTGCCCG	4074
QY	4290	GCTCAGCTGTGTGCTCCCTCCGCCCCAGGACAGTGAACCGGCTCTG	GGGGAAGTGGTTACG	4349
DB	4075	GCTCAGCTGTGTGCTCCCTCCGCCCCAGGACAGTGAACCGGCTCTG	GGGGAAGTGGTTACG	4134
QY	4350	AACCTGACAGGCGCGGAAACCTGTCTGACTTCTCGGTCAAGACCT	TACCGCGGCTGTGGTCGCG	4409

Db	4135		AACCTGACAGGCGCGAACCTGTCTGACTCTCTCGTCAAGACCTTACCCGCGCCCTGGTGCGC	4194
Qy	4410	CAGGGCCTGAAGACTAAAGAAATGGGTCAAATGAGGTACAGGTACGAGAGCTTCTCGCTGGGG	4469	
Db	4195	CAGGGCCTGAAGACTAAGAAATGGGTGAATGAGGTACAGATACGAGAGCTTCTCGCTGGGG	4254	
Qy	4470	GGCCGAGACCAAGCCTCGCCCTCGGGCCAAAGAGTGTGGGCCGCTCAGTGGAGGAGTTGTGG	4529	
Db	4255	GGCCGAGACCAAGCCTCGCCCTCGGGCCAAAGAGTGTGGGCCGCTCAGTGGAGGAGTTGTGG	4314	
Qy	4530	CGCGTGTGAGTCCCTCGCTCGGCGGGGCCCTCGACCGTGTCTCTGAAAAACCTCACAGCC	4589	
Db	4315	CGCGTGTGAGTCCCTCGCTCGGCGGGGCCCTCGACCGTGTCTCTGAAAAACCTCACAGCC	4374	
Qy	4590	TGGGCTCACAGCCTGGATGCTCAGGACAGATCTCAAGATCTGGTTCAACAACAAGGCTGG	4649	
Db	4375	TGGGCTCACAGCCTGGATGCTCAGGACAGATCTCAAGATCTGGTTCAACAACAAGGCTGG	4434	
Qy	4650	CACTCCAATGGTGGCTTTGTCAACCGAGCGAGCAACGCAATCTCCGTGCTCACCTGCCCC	4709	
Db	4435	CACTCCAATGGTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCCGTGCTCACCTGCCCC	4494	
Qy	4710	CCAGGCCGGCCCGGCGACGGCCACAGCATCACCACTCAACACCCCTTGAACCTCACCC	4769	
Db	4495	CCAGGCCGGCCCGGCGACGGCCACAGCATCACCACTCAACACCCCTTGAACCTCACCC	4554	
Qy	4770	AAGGAGCAGCTGCTGAGGGCTGCACATGAGCCCTCTCGGTGACGCTCCGCTCCATC	4829	
Db	4555	AAGGAGCAGCTGCTGAGGGCTGCACATGAGCCCTCTCGGTGACGCTCCGCTCCATC	4614	
Qy	4830	TGTTGTGCTTTTGCCATGTCTTTGTGCCGGCCAGCTTCACTCTGTCTCATTTGAGGAG	4889	
Db	4615	TGTTGTGCTTTTGCCATGTCTTTGTGCCGGCCAGCTTCACTCTGTCTCATTTGAGGAG	4674	
Qy	4890	CGAGTCAACCGAGCCAAAGCACTGCAAGCTCATGGGGGGGCTGTCCCCCAACCTCTACTGG	4949	
Db	4675	CGAGTCAACCGAGCCAAAGCACTGCAAGCTCATGGGGGGGCTGTCCCCCAACCTCTACTGG	4734	
Qy	4950	CTTGGCAACTTTCTCTGGGACATGTGTACTCTGTGGTGCCACGATGCTCGTGGTGCTC	5009	
Db	4735	CTTGGCAACTTTCTCTGGGACATGTGTACTCTGTGGTGCCACGATGCTCGTGGTGCTC	4794	
Qy	5010	ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTCTGCCAACTGTCTCTCTCCCTG	5069	
Db	4795	ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTCTGCCAACTGTCTCTCTCCCTG	4854	
Qy	5070	CTGTTGTCTACTACTGATATGCTGGTCATACACCGCTCATGTAPCCAGCCTCTTCTTTC	5129	
Db	4855	CTGTTGTCTACTACTGATATGCTGGTCATACACCGCTCATGTAPCCAGCCTCTTCTTTC	4914	
Qy	5130	TTCTCCGTGCCAGCAGCCTATGTGGTCTCACCTGCTATAAACCTCTTTATTGGCATC	5189	
Db	4915	TTCTCCGTGCCAGCAGCCTATGTGGTCTCACCTGCTATAAACCTCTTTATTGGCATC	4974	
Qy	5190	AATGGAAGCATGCGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAAGCTGACGAGGTG	5249	
Db	4975	AATGGAAGCATGCGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAAGCTGACGAGGTG	5034	
Qy	5250	AGCCGGATCTGAAACAGGTCCTTCTTATCTTCCCACTTTGTGCTTGGCCCGGGGGCTC	5309	
Db	5035	AGCCGGATCTGAAACAGGTCCTTCTTATCTTCCCACTTTGTGCTTGGCCCGGGGGCTC	5094	
Qy	5310	ATTGACATGTTGGGAACAGGGCCATGGCTGATGCCCTTTGAGGCTTTGGGACAGAGGCAG	5369	
Db	5095	ATTGACATGTTGGGAACAGGGCCATGGCTGATGCCCTTTGAGGCTTTGGGACAGAGGCAG	5154	
Qy	5370	TTCCAGTCACTCCGTGCGCTGGGAGGTGGTCGGAAGAACTCTTGGCCATGTGTATACAG	5429	
Db	5155	TTCCAGTCACTCCGTGCGCTGGGAGGTGGTCGGAAGAACTCTTGGCCATGTGTATACAG	5214	
Qy	5430	GGGCCCCCTTCTCTTTCATCACTACTGCTGCAGCACCGAAGCCAACTCTCTGCCACAG	5489	

5215	GGGCGCCTCTTCTCTTTCACACTACTCTCTGCGACACCGAAGCCAACTCTTCGCCACAG	5274
QY	CCCAGGCTGAGGTCTCTTGCCACTCTTGCGGAGAGGAGACGAGGATGTAGCCCTGAACCG	5549
	5275	CCCAGGCTGAGGTCTCTTGCCACTCTTGCGGAGAGGAGCGAGGATGTAGCCCTGAACCG
QY	GAGCGGCTGTCCAAGGAGCCACCAGGGGGATGTGTTGTGTGTGAGGAACTTTGACCAAG	5609
	5335	GAGCGGCTGTCCAAGGAGCCACCAGGGGGATGTGTTGTGTGTGAGGAACTTTGACCAAG
QY	GTATACGTTGGGCAGAGGATCCAGCTGTATTGACCGCTGTGCTCGGGATTCCTCCCTGGT	5669
	5395	GTATACGTTGGGCAGAGGATTCAGCTGTATTGACCGCTGTGCTCGGGATTCCTCCCTGGT
QY	GAGTGTTTTGGGCTGCTGGGTGTGAATGGAGCAGGGAGAAGCACTCCAGTTTCGATGGTG	5729
	5455	GAGTGTTTTGGGCTGCTGGGTGTGAATGGAGCAGGGAGAAGCACTCCAGTTTCGATGGTG
QY	ACGGGGACACATTTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACACGCTGGCCCGG	5789
	5515	ACGGGGACACATTTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACACGCTGGCCCGG
QY	GAACCCAGTCTCGGCACCTCAGCATGGGATCTGCCCTCAATCCGATGGCATCTTTGAG	5849
	5575	GAACCCAGTCTCGGCACCTCAGCATGGGATCTGCCCTCAATCCGATGGCATCTTTGAG
QY	CTGCTGACGGGCGCGAGCACCTTGGAGCTGCTTTGGCGGCTCTGCGGTGTCCCGAGGCC	5909
	5635	CTGCTGACGGGCGCGAGCACCTTGGAGCTGCTTTGGCGGCTCTGCGGTGTCCCGAGGCC
QY	CAGTTGCCAGAACCGGTGGCTCAGGCGCTGGCGGCTCTGSGACTCTCATGGTACGCAGAC	5969
	5695	CAGTTGCCAGAACCGGTGGCTCAGGCGCTGGCGGCTCTGSGACTCTCATGGTACGCAGAC
QY	CGGCTGCGAGGACCTACAGCGAGGAGAA CA AACAGCAAGCTGCGGACCGGCCCTGGCGCTG	6029
	5755	CGGCTGCGAGGACCTACAGCGAGGAGAA CA AACCGAAGCTGCGGACCGGCCCTGGCGCTG
QY	GTTTGGGACCCAGCCGCTGGTCTTGGACGAGCCGACACAGGCA TGGACCCAGCGCG	6089
	5815	GTTTGGGACCCAGCCGCTGGTCTTGGACGAGCCGACACAGGCA TGGACCCAGCGCG
QY	CGGCGCTTCTTTTGGAA CAGCCCTTTTGGCCGCTGGTGGGAGGGCCGTTCAGTGATGCTC	6149
	5875	CGGCGCTTCTTTTGGAA CAGCCCTTTTGGCCGCTGGTGGGAGGGCCGTTCAGTGATGCTC
QY	ACCTCCCATAGCATGGAGGAGTGTAAGCCCTCTGCTCGGCCCTAGCCATCATGTTGTAAT	6209
	5935	ACCTCCCATAGCATGGAGGAGTGTAAGCCCTCTGCTCGGCCCTAGCCATCATGTTGTAAT
QY	GGCGGTTCCGCTGCGCCGACGCCCAACATCTCAAGGGCAGATTCGCGGGGGGTTCAC	6269
	5995	GGCGGTTCCGCTGCGCCGACGCCCAACATCTCAAGGGCAGATTCGCGGGGGGTTCAC
QY	ACACTGACCTTGGGGTGCCCGCGCNAAGSTCCACGCCGACAGGCATTCGCGGGGGGTTCAC	6329
	6055	ACACTGACCTTGGGGTGCCCGCGCNAAGSTCCACGCCGACAGGCATTCGCGGGGGGTTCAC
QY	GAGTTCCTCGGTCGAGGCTGCGCGAGGCA CATGAGGCGCGCTTCAGCTGCGCG	6389
	6115	GAGTTCCTCGGTCGAGGCTGCGCGAGGCA CATGAGGCGCGCTTCAGCTGCGCG
QY	CCGGGAGGGCGCTGCGGCCCTTTTGGAGAGCTGGCGGTGCA CGGGCGCAG	6449
	6175	CCGGGAGGGCGCTGCGGCCCTTTTGGAGAGCTGGCGGTGCA CGGGCGCAG
QY	CACGGCTGAGGACTTTTCCGTGAGCCAGCATGCTGGAGGAGGTATTTCTTGACTTC	6509
	6235	CACGGCTGAGGACTTTTCCGTGAGCCAGCATGCTGGAGGAGGTATTTCTTGACTTC
QY	TCAAAGGACAGGGGAA GGA CGAGGACACCCGAA GAGCAGAGGAGGCGAGGTGGGAGTG	6569
	6295	TCAAAGGACAGGGGAA GGA CGAGGACACCCGAA GAGCAGAGGAGGCGAGGTGGGAGTG

Qy	6570	GA	CCCCCGCCAGCGCTG	CAGCA	CCCCAAACGGT	CAGCCAGT	TCTCGATG	ACCTTAGC	6629																	
Db	6355	GA	CCCCCGCCAGCGCTG	CAGCA	CCCCAAACGGT	CAGCCAGT	TCTCGATG	ACCTTAGC	6414																	
Qy	6630	ACTGCCGAGACTGTG	CTC	6647																						
Db	6415	ACTGCCGAGACTGTG	CTC	6432																						
RESULT 11																										
US-10-154-419-3																										
; Sequence 3, Application US/10154419																										
; Publication NO. US20030143675A1																										
; GENERAL INFORMATION:																										
; APPLICANT: Curtis, Rory A.J.																										
; APPLICANT: Gluckeman, Maria Alexandra																										
; APPLICANT: Meyers, Rachel E.																										
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,																										
; TITLE OF INVENTION: 57259, 67118, 67067, 8099, 46455, 54414, 53763,																										
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 572,																										
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR																										
; FILE REFERENCE: MNI-249																										
; CURRENT APPLICATION NUMBER: US/10/154,419																										
; CURRENT FILING DATE: 2002-05-22																										
; Prior Application removed - See Palm or File Wrapper																										
; NUMBER OF SEQ ID NOS: 99																										
; SOFTWARE: FastSEQ Version 4.0																										
; SEQ ID NO 3																										
; LENGTH: 6432																										
; TYPE: DNA																										
; ORGANISM: Homo sapiens																										
; FEATURE:																										
; NAME/KEY: CDS																										
; LOCATION: (1)..(6432)																										
US-10-154-419-3																										
Query Match 94.0%; Score 6398.4; DB 15; Length 6432;																										
Best Local Similarity 99.7%; Pred. No. 0;																										
Matches 6421; Conservative 0; Mismatches 11; Indels 6; Gaps 1;																										
Qy	210	ATG	GCCTTC	TGG	AC	AGCTG	ATGCTG	CTCTG	GAAGAAATTT	CATG	TATCC	CGG	GAGA	269												
Db	1	ATG	GCCTTC	TGG	AC	AGCTG	ATGCTG	CTCTG	TGG	AAATTT	CATG	TATCC	CGG	GAGA	60											
Qy	270	CAG	CGGT	CAG	CTC	TCTG	GTG	CGA	ATTTG	CTG	CGCT	CTCT	CTCT	CTT	CA	TCTT	CA	TCTT	GGT	329						
Db	61	CAG	CGT	-----	GT	CTCT	GTG	CGA	ATTTG	CTG	CGCT	CTCT	CTCT	CTT	CA	TCTT	CA	TCTT	GGT	114						
Qy	330	GCT	GTTC	CG	CACT	CCCA	CGCC	CCCT	GG	AGC	ACAT	GAAT	GC	CACT	TCC	CAAA	CAAG	CCCA	389							
Db	115	GCT	GTTC	CG	CACT	CCCA	CGCC	CCCT	GG	AGC	ACAT	GAAT	GC	CACT	TCC	CAAA	CAAG	CCCA	174							
Qy	390	CTG	CCAT	CG	CG	GG	CA	CCG	TG	CTC	CA	GGG	TCT	CA	TG	TG	TAT	GT	GA	CA	CACC	449				
Db	175	CTG	CCAT	CG	CG	GG	CA	CCG	TG	CTC	CA	GGG	TCT	CA	TG	TG	TAT	GT	GA	CA	CACC	234				
Qy	450	TG	CTTT	CCG	AG	CTG	AC	CGG	CG	AGG	AG	CCG	CG	CG	CTG	AG	CA	CTT	CA	AG	CT	CC	509			
Db	235	TG	CTTT	CCG	AG	CTG	AC	CGG	CG	AGG	AG	CCG	CG	CG	CTG	AG	CA	CTT	CA	AG	CT	CC	294			
Qy	510	CTG	GTCT	CC	GG	TG	CTAG	CCG	ATG	CCG	CA	CTG	TG	CTG	G	AG	GGG	CG	CAG	TG	CC	CAGG	569			
Db	295	CTG	GTCT	CC	GG	TG	CTAG	CCG	ATG	CCG	CA	CTG	TG	CTG	T	G	AG	GGG	CG	CAG	TG	CC	CAGG	354		
Qy	570	ACG	CTG	CT	GG	CC	TAG	GA	AG	CTG	AT	CG	CA	C	CT	G	AG	GG	CT	CG	AG	CA	CGG	CC	CAG	629
Db	355	ACG	CTG	CT	GG	CC	TAG	GA	AG	CTG	AT	CG	CA	C	CT	G	AG	GG	CT	CG	AG	CA	CGG	CC	CAG	414
Qy	630	CCT	CA	AC	CA	CC	AG	C	AG	TG	CT	CA	CT	GG	AA	CC	AC	CC	AT	TG	CT	GG	AG	CT	CTG	689
Db	415	CCT	CA	AC	CA	CC	AG	C	AG	TG	CT	CA	CT	GG	AA	CC	AC	CC	AT	TG	CT	GG	AG	CT	CTG	474

QY 690 ACGTCACTGCTCGGACGGAATCCCTGGGGTTGGCACTGGGCCAAGCCAGAGCCCTTG 749
Db 475 ACGTCACTGCTCGGACGGAATCCCTGGGGTTGGCACTGGGCCAAGCCAGAGCCCTTG 534
QY 750 CACAGCTTTGGAGCCGCTGAGGACCTGGCCAGGAGCTCTCGGCGCTGCGGACCTG 809
Db 535 CACAGCTTTGGAGCCGCTGAGGACCTGGCCAGGAGCTCTCGGCGCTGCGGACCTG 594
QY 810 GTGGAGCTTTGGGCACTGCTGCAGAGACCCCGAGGAGCCAGCGGCCCTCGAGTTGCTG 869
Db 595 GTGGAGCTTTGGGCACTGCTGCAGAGACCCCGAGGAGCCAGCGGCCCTCGAGTTGCTG 654
QY 870 TCAGAGGCCCTCTGAGTGTGAGGGGACCTAGCAGACAGTGGGGCCCTCCCTCAACTGG 929
Db 655 TCAGAGGCCCTCTGAGTGTGAGGGGACCTAGCAGACAGTGGGGCCCTCCCTCAACTGG 714
QY 930 TACGAGGCTAGTGACCTGATGAGAGCTGTGGGGCAGGAGCCAGAAATCGCCCTGCCAGC 989
Db 715 TACGAGGCTAGTGACCTGATGAGAGCTGTGGGGCAGGAGCCAGAAATCGCCCTGCCAGC 774
QY 990 AGCAGCCTGAGCCCGCCCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCCGCTGTCC 1049
Db 775 AGCAGCCTGAGCCCGCCCTGCTCGGAGCTGATTTGGAGCCCTGGACAGCCACCCCGCTGTCC 834
QY 1050 GGCCTGCTCTGAGAGCGCCTGAAGCCTCTGATCTCTGGGAAAGCTACTCTTTGACACAGAT 1109
Db 835 GGCCTGCTCTGAGAGCGCCTGAAGCCTCTGATCTCTGGGAAAGCTACTCTTTGACACAGAT 894
QY 1110 ACACCTTTTACCCGGAAGCTATGCCCCAGGTGAACCCGACCTTCGAGGAGCTACCCCTG 1169
Db 895 ACACCTTTTACCCGGAAGCTATGCCCCAGGTGAACCCGACCTTCGAGGAGCTACCCCTG 954
QY 1170 CTGAGGAGTGTCCGGAGGTGTGGAGATGCTTGGGACCCCGGATCTTCACTTCAATGAAC 1229
Db 955 CTGAGGAGTGTCCGGAGGTGTGGAGATGCTTGGGACCCCGGATCTTCACTTCAATGAAC 1014
QY 1230 GACAGTTCCAAATGTGGCCATGCTCAGCGGCTCTGACGAGTGCAGATGAAGGAAGAG 1289
Db 1015 GACAGTTCCAAATGTGGCCATGCTCAGCGGCTCTGACGAGTGCAGATGAAGGAAGAG 1074
QY 1290 CAGCCAGACCTGGAGGCGGGACCAATGAGGCGCTGGATCTTTCTGGAACCTTGGG 1349
Db 1075 CAGCCAGACCTGGAGGCGGGACCAATGAGGCGCTGGATCTTTCTGGAACCTTGGG 1134
QY 1350 AGCGTGTGCTACAGCTGCGAGGACGCAACGCTGATGTGGGCACTTGGTGGGACGCTG 1409
Db 1135 AGCGTGTGCTACAGCTGCGAGGACGCAACGCTGATGTGGGCACTTGGTGGGACGCTG 1194
QY 1410 GGCCGAGTGAAGGAGTGTCTGCTTGGACAAAGCTGGAGGCGGCACTTCAAGAGCAGCC 1469
Db 1195 GGCCGAGTGAAGGAGTGTCTGCTTGGACAAAGCTGGAGGCGGCACTTCAAGAGCAGCC 1254
QY 1470 CTGGTGTGCGGGCCCTGCAACTGCTCGGGAACATCGATTTGGGCGGGCTGCTTTC 1529
Db 1255 CTGGTGTGCGGGCCCTGCAACTGCTCGGGAACATCGATTTGGGCGGGCTGCTTTC 1314
QY 1530 TTGGGACCTGAGGACTCTTTACAGACCCCAAGAGCACCACCCAGACCTTGGGCCCGGC 1589
Db 1315 TTGGGACCTGAGGACTCTTTACAGACCCCAAGAGCACCACCCAGACCTTGGGCCCGGC 1374
QY 1590 CAGGTGCGCATCAAAATCCGATGACATTTGACGTTGCAAGGACCAATAAGATCAGG 1649
Db 1375 CAGGTGCGCATCAAAATCCGATGACATTTGACGTTGTTGACGTTGTTGACGTTGTTG 1434
QY 1650 GACAGGTTTGGGACCTTGGCCAGCCCGGAGCCCTTGAACCGGACCTTGGCTACGTGTG 1709
Db 1435 GACAGGTTTGGGACCTTGGCCAGCCCGGAGCCCTTGAACCGGACCTTGGCTACGTGTG 1494
QY 1710 GCGGCTTGTGTACTGCTCAAGACCTGTGTGAGCGTGCAGCGCTCGGCTGTCTAGCGGC 1769
Db 1495 GCGGCTTGTGTACTGCTCAAGACCTGTGTGAGCGTGCAGCGCTCGGCTGTCTAGCGGC 1554
QY 1770 GCCAAACCCCGGCGGCTCTACTCTGACGAGATGCGCTATCTCGTGTATGTGACGAC 1829

Db 1555 GCCAAACCCCGGCGGCTCTACTCTGACGAGATGCGCTATCCGTGCTATCTGGACGAC 1614
QY 1830 GTGTTCCTGCTGCTGAGCGGCTGCTGCGGCTTCTTCTGACGCTGCGCTGATCTAC 1889
Db 1615 GTGTTCCTGCTGCTGAGCGGCTGCTGCGGCTTCTTCTGACGCTGCGCTGATCTAC 1674
QY 1890 TCCGTGACACTGACAGTGAAGGCCGTGTGCGGGAGAGGAGACGCGGTGCGGACACC 1949
Db 1675 TCCGTGACACTGACAGTGAAGGCCGTGTGCGGGAGAGGAGACGCGGTGCGGACACC 1734
QY 1950 ATGCGGCGCATGCGGCTCAGCGCGCGGCTGCTGCGCTAGGCTGCTTCTGAGCTGCTC 2009
Db 1735 ATGCGGCGCATGCGGCTCAGCGCGCGGCTGCTGCGCTAGGCTGCTTCTGAGCTGCTC 1794
QY 2010 GGGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTTCTGCTCAAGCTGCGGGACATCTCT 2069
Db 1795 GGGCCCTTCTGCTCAGCGCGGCTGCTGCTGCTTCTGCTCAAGCTGCGGGACATCTCT 1854
QY 2070 CCCTACAGCCACCCGGGCGTGTCTTCTGTTTCTGTCGAGCCTTCTCGGCTGCGGCTG 2129
Db 1855 CCCTACAGCCACCCGGGCGTGTCTTCTGTTTCTGTCGAGCCTTCTCGGCTGCGGCTG 1914
QY 2130 ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGGCAACCTTGGCTGCGGCTGCGG 2189
Db 1915 ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCAACCTTGGCTGCGGCTGCGG 1974
QY 2190 GGCCTGGGCTACTTCTCTCTCTACCTGCTTCTGCTGCTGCTGCTTGGCTGCGGACCGG 2249
Db 1975 GGCCTGGGCTACTTCTCTCTCTCTACCTGCTTCTGCTGCTGCTGCTTGGCTGCGGACCGG 2034
QY 2250 CTGCCCGCGGGTGGCGCGTGGCGGAGCCTGCTGCTGCTGCTGCTGCTTGGCTTGGGCT 2309
Db 2035 CTGCCCGCGGGTGGCGCGTGGCGGAGCCTGCTGCTGCTGCTGCTGCTTGGCTTGGGCT 2094
QY 2310 TGGAGAGCCTTGGCTCTGCTGAGAGCAGGCGGAGGCGGCGCACTGCGCACTGCTGCGG 2369
Db 2095 TGGAGAGCCTTGGCTCTGCTGAGAGCAGGCGGAGGCGGCGCACTGCGCACTGCTGCGG 2154
QY 2370 ACCCGGCTACGCGAGAGCTCTTCTGAGCCTGCGGCGGAGGCTCTGCTGCTGCTGAG 2429
Db 2155 ACCCGGCTACGCGAGAGCTCTTCTGAGCCTGCGGCGGAGGCTCTGCTGCTGCTGAG 2214
QY 2430 GCGGCGCTTACGCGCTCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
Db 2215 GCGGCGCTTACGCGCTCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
QY 2490 ATCCCTGAACCAATGGAATTTTCTTCTGAGGAGCTACTGCTGCGGACCTTGGGCTG 2549
Db 2275 ATCCCTGAACCAATGGAATTTTCTTCTGAGGAGCTACTGCTGCGGACCTTGGGCTG 2334
QY 2550 AAGAGTCCAGCCCTTGGCCCAACCCGCTGGACCCAAAGGCTGCTGCTGCTGCTGCTG 2609
Db 2335 AAGAGTCCAGCCCTTGGCCCAACCCGCTGGACCCAAAGGCTGCTGCTGCTGCTGCTG 2394
QY 2610 CCGGCTGAGTCTGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2669
Db 2395 CCGGCTGAGTCTGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2454
QY 2670 CAGCGAGCCCTTGGGCGGCTCAGCTGAGCTTCTTCTGAGGCGCACTGCTGCTGCTG 2729
Db 2455 CAGCGAGCCCTTGGGCGGCTCAGCTGAGCTTCTTCTGAGGCGCACTGCTGCTGCTGCTG 2514
QY 2730 GGGCACAACCGGGCGGCAAGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2789
Db 2515 GGGCACAACCGGGCGGCAAGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2574
QY 2790 AGTGTGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2849
Db 2575 AGTGTGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2634
QY 2850 CCCACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2909

Db 2635 CCCACCTGGCGCTCTGTCTCTCAGTACAA CGTGTCTTTTGACATGCTGACCGTGGACGAG 2694
Qy 2910 CAGCTCTGGTTCTATGGGCGGCTGAAGGGTCTGAGTCCCGCTGTAGTGGGCCCCGAGCAG 2969
Db 2695 CAGCTCTGGTTCTATGGGCGGCTGAAGGGTCTGAGTCCCGCTGTAGTGGGCCCCGAGCAG 2754
Qy 2970 GACCGTCTGCTGAGGATGTGGGCTGGTCTCCAAAGCAGAGTGTGAGACTTCGCCACCTC 3029
Db 2755 GACCGTCTACTGCAAGGATGTGGGCTGGTCTCCAAAGCAGAGTGTGAGACTTCGCCACCTC 2814
Qy 3030 TCTGGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 3089
Db 2815 TCTGGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCCTTTGTGGGCGGCTCCCAAGTT 2874
Qy 3090 GTTATCTCGACGAGCTTACGGCTGGCTGGATCTCTGCTTCCCGCCCGCGGTATTTGGGAG 3149
Db 2875 GTTATCTCGACGAGCTTACGGCTGGCTGGATCTCTGCTTCCCGCCCGCGGTATTTGGGAG 2934
Qy 3150 CTGCTGCTCAATACCGAGAGGTCCGACGCTGATCTCTCCACCCACACCTGGATGAG 3209
Db 2935 CTGCTGCTCAATACCGAGAGGTCCGACGCTGATCTCTCCACCCACACCTGGATGAG 2994
Qy 3210 GCAGAGCTGTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCGCTGTGCTGTGGC 3269
Db 2995 GCAGAGCTGTGGGAGACCGTGTGGCGTGGTGGCAGGTGGCGCTGTGCTGTGGC 3054
Qy 3270 TCCCACTCTTCTTGCGCCGCTCACTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC 3329
Db 3055 TCCCACTCTTCTTGCGCCGCTCACTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC 3114
Qy 3330 CGCTGCCCCGTACCCCAATGAGAGGCTGCACTGACATGAGGGCAGTGTGGACCC 3389
Db 3115 CGCTGCCCCGTACCCCAATGAGAGGCTGCACTGACATGAGGGCAGTGTGGACCC 3174
Qy 3390 AGGCAGAAAAGAAATGGCAGCCAGGGCAGCAGAGTCCGCACTCTCAGCTGCTGGCC 3449
Db 3175 AGGCAGAAAAGAAATGGCAGCCAGGGCAGCAGAGTCCGCACTCTCAGCTGCTGGCC 3234
Qy 3450 CTGGTACAGCACTGGGTGCCCCGGGCAACGGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3509
Db 3235 CTGGTACAGCACTGGGTGCCCCGGGCAACGGCTGGTGGAGAGCTGCCACACGAGCTGGTG 3294
Qy 3510 CTGGTGTCCCTTACACGGGTGCCATGACGGGAGCTCCGCACTCTCAGCTGCTGGCC 3569
Db 3295 CTGGTGTCCCTTACACGGGTGCCATGACGGGAGCTTCCGCACTCTTCCGAGAGCTA 3354
Qy 3570 GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAG 3629
Db 3355 GACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAG 3414
Qy 3630 GAGATCTTCTGAAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGC 3689
Db 3415 GAGATCTTCTGAAAGTGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGC 3474
Qy 3690 TGGGGGAGCATTATGACAGGATTTGCTGGCTAGACCTTACCTGCGGCTCAAGATG 3749
Db 3475 TGGGGGAGCATTATGACAGGATTTGCTGGCTAGACCTTACCTTACGGCTCAAGATG 3534
Qy 3750 CCGCACAGGAGACAGCGCTGGAGAACGGGAACAGCTGGGTGAGCGCCACAGACTGAC 3809
Db 3535 CCGCACAGGAGACAGCGCTGGAGAACGGGAACAGCTGGGTGAGCGCCACAGACTGAC 3594
Qy 3810 CAGGGCTCTGGGCGCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCGCAGCAG 3869
Db 3595 CAGGGCTCTGGGCGCAGACGCGTGGGCGGGTACAGGGCTGGGCACTGACCCGCGCAGCAG 3654
Qy 3870 CTCAGGCCCTGCTTCTCAAGCGCTTCTGCTGGCGCGCGCAGCGCGCGGCGCTGTC 3929
Db 3655 CTCAGGCCCTGCTTCTCAAGCGCTTCTGCTGGCGCGCGCAGCGCGCGGCGCTGTC 3714
Qy 3930 GCCAGATCGTGTGCTGCGCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTG 3989
Db 3715 GCCAGATCGTGTGCTGCGCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCGTG 3774

Qy 3990 CCTCTTTTCGGGCACTAACCGGCTCTCGGCTCAGTCCACCATGTACGGTGTCCAGTG 4049
Db 3775 CCTCTTTTCGGGCACTAACCGGCTCTCGGCTCAGTCCACCATGTACGGTGTCCAGTG 3834
Qy 4050 TCCTTCTTTCAGTGAAGACGCCCCAGGGGACCTTGACGCTGCCCGCTGTCTGAGGCGCTG 4109
Db 3835 TCCTTCTTTCAGTGAAGACGCCCCAGGGGACCTTGACGCTGCCCGCTGTCTGAGGCGCTG 3894
Qy 4110 CTGACGAGGCAAGGACTGAGGAGGAGCCCCCAGTGCAGCATAGTCTCCACAGGTTCTCGGCA 4169
Db 3895 CTGACGAGGCAAGGACTGAGGAGGAGCCCCCAGTGCAGCATAGTCTCCACAGGTTCTCGGCA 3954
Qy 4170 CGAAGATTTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAATCGGACCCAGAGTCT 4229
Db 3955 CGAAGATTTCTGCTGAAGTGGCCAAAGTCTTGGCCAGTGGCAATCGGACCCAGAGTCT 4014
Qy 4230 CCATCCCGAGCCTGCAGTGTAGCAGCCGGTGCCTGGCGGCTGTGCGCCGACTGCGCG 4289
Db 4015 CCATCCCGAGCCTGCAGTGTAGCAGCCGGTGCCTGGCGGCTGTGCGCCGACTGCGCG 4074
Qy 4290 GCTGACGCTGGTGGTCCCTCCCGCCCCAGGCAAGTGCAGGCTCTGGGGAAGTGGTTGAG 4349
Db 4075 GCTGACGCTGGTGGTCCCTCCCGCCCCAGGCAAGTGCAGGCTCTGGGGAAGTGGTTGAG 4134
Qy 4350 AACCTGACAGGCGCGGAACTGTCTGACTTCTGCTCAAGACCTACCGCGCCTGGTGCGC 4409
Db 4135 AACCTGACAGGCGCGGAACTGTCTGACTTCTGCTCAAGACCTACCGCGCCTGGTGCGC 4194
Qy 4410 CAGGGCTCAAGACTGAAGTGGGTGAATGAGGTACGGTACGGAGCTTCTCGCTGGGG 4469
Db 4195 CAGGGCTCAAGACTGAAGTGGGTGAATGAGGTACGGTACGGAGGCTTCTCGCTGGGG 4354
Qy 4470 GSCCGAGACCCAGGCTGCGCTCCCGCCCAAGAGTTGGGCGCTCACTGAGAGGTTGTGG 4529
Db 4255 GSCCGAGACCCAGGCTGCGCTCCCGCCCAAGAGTTGGGCGCTCACTGAGAGGTTGTGG 4314
Qy 4530 GCGCTCTGAGTCCCTGCTGCGGGGCGCTCGACGGTGTCTGAAAACCTCACAGCC 4589
Db 4315 GCGCTCTGAGTCCCTGCTGCGGGGCGCTCGACGGTGTCTGAAAACCTCACAGCC 4374
Qy 4590 TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGG 4649
Db 4375 TGGGCTCACAGCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCAACAAACAAAGGCTGG 4434
Qy 4650 CACTCCATGGTGGCTTTGTCAA CCGAGCCAGCAA CGCAATCTCCGTGTCTACCTGCCC 4709
Db 4435 CACTCCATGGTGGCTTTGTCAA CCGAGCCAGCAA CGCAATCTCCGTGTCTACCTGCCC 4494
Qy 4710 CAGGGCGCGGCGCCAGCCGACGATCACACACTCAA CCAACCCCTTGAACCTCACCC 4769
Db 4495 CAGGGCGCGGCGCCAGCCGACGATCACACACTCAA CCAACCCCTTGAACCTCACCC 4554
Qy 4770 AAGGAGCAGCTGTCTGAGGCTGACATGAGGCTCTCGGTGAGACGCTCTGCTCCATC 4829
Db 4555 AAGGAGCAGCTGTCTGAGGCTGACATGAGGCTCTCGGTGAGACGCTCTGCTCCATC 4614
Qy 4830 TGTGTGGTCTTTGCCATGTCTTTGTCCCGGCGAGCTTCACTCTTGTCTCATTTGAGGAG 4889
Db 4615 TGTGTGGTCTTTGCCATGTCTTTGTCCCGGCGAGCTTCACTCTTGTCTCATTTGAGGAG 4674
Qy 4890 CGAGTCAACCGAGCCAGCAACCTGACGATCATGCGGGGCGCTGCCCCACCTTACTGCG 4949
Db 4675 CGAGTCAACCGAGCCAGCAACCTGACGATCATGCGGGGCGCTGCCCCACCTTACTGCG 4734
Qy 4950 CTTTGGCAACTTTCTCTGSGACATGTGTAACTACTTGTGCGCAGCATGTCATGCTGCTC 5009
Db 4735 CTTTGGCAACTTTCTCTGSGACATGTGTAACTACTTGTGCGCAGCATGTCATGCTGCTC 4794
Qy 5010 ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTCGCCAA CTTGCTGTCTCTCTG 5069
Db 4795 ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCTCGCCAA CTTGCTGTCTCTCTG 4854

QY 5070 CTGTTGCTACTCTGCTATGGCTGTGATCACAACGGCTCATGTACCCAGCCTCTCTTTC 5129
Db 4855 CTGTTGCTACTACTGATGCTGTGTCGATCACACCGCTCATGTACCCAGCCTCTCTTTC 4914
QY 5130 TTCTCCGTCGCCAGCACAGCCTATGTGTGCTCACTCATATAAACCCTTTATTGGCATC 5189
Db 4915 TTCTCCGTCGCCAGCACAGCCTATGTGTGCTCACTCATATAAACCCTTTATTGGCATC 4974
QY 5190 AATGGAACATGGCCACCTTTGTGCTTGAGCTCTTCTGTATCAAGAGCTGCAGAGGTG 5249
Db 4975 AATGGAACATGGCCACCTTTGTGCTTGAGCTCTTCTGTATCAAGAGCTGCAGAGGTG 5034
QY 5250 AGCCGGATCTTGAACAAGCTTTCTTATCTTCCCCCACTTCTGTGTCGGGCGGGGGCTC 5309
Db 5035 AGCCGGATCTTGAACAAGCTTTCTTATCTTCCCCCACTTCTGTGTCGGGCGGGGGCTC 5094
QY 5310 ATTGACATGGTCGGAACACAGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAG 5369
Db 5095 ATTGACATGGTCGGAACACAGCCATGCTGATGCTTTGAGCGCTTGGGAGACAGGCAG 5154
QY 5370 TTCCAGTCAACCCCTCGCTGGGAGGTGTCGCAAGAACCTCTTGGCCCATGGTATACAG 5429
Db 5155 TTCCAGTCAACCCCTCGCTGGGAGGTGTCGCAAGAACCTCTTGGCCCATGGTATACAG 5214
QY 5430 GGGCCCCCTCTTCTTCTTCACTACTGTGACGACCGAAGCCAACTCTCTGCAACAG 5489
Db 5215 GGGCCCCCTCTTCTTCTTCACTACTGTGACGACCGAAGCCAACTCTCTGCAACAG 5274
QY 5490 CCCAGGCTGAGGTCTCTGCCACTCTGGGAGGAGGACGAGGATGTAGCCGCTGAACGG 5549
Db 5275 CCCAGGCTGAGGTCTCTGCCACTCTGGGAGGAGGACGAGGATGTAGCCGCTGAACGG 5334
QY 5550 GAGCGGCTGGTCCAAAGAGCCACCCAGCGGATGTGTTGGTCTGAGGAACTTGACCAAG 5609
Db 5335 GAGCGGCTGGTCCAAAGAGCCACCCAGCGGATGTGTTGGTCTGAGGAACTTGACCAAG 5394
QY 5610 GTATACCGTGGGACAGGATGCCAGCTGTGACCGCTTGTGCTGGGATTCCTCCCTGGT 5669
Db 5395 GTATACCGTGGGACAGGATGCCAGCTGTGACCGCTTGTGCTGGGATTCCTCCCTGGT 5454
QY 5670 GAGTGTGTTGGGCTGCTGGTGTGAAATGAGGAGGAAAGCTCAGCTTTGCGATGTTG 5729
Db 5455 GAGTGTGTTGGGCTGCTGGTGTGAAATGAGGAGGAAAGCTCAGCTTTGCGATGTTG 5514
QY 5730 ACGGGGACATATGGCCAGACGAGGCGAGGCTGTGCTGGAGGCGACAGAGTGGCCCGG 5789
Db 5515 ACGGGGACATATGGCCAGACGAGGCGAGGCTGTGCTGGAGGCGACAGAGTGGCCCGG 5574
QY 5790 GAAACCCAGTGTGCGCACTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAG 5849
Db 5575 GAAACCCAGTGTGCGCACTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAG 5634
QY 5850 CTGCTGACGGGCGGCGACACTGTGAGCTGCTTGGCGCGCTGCGCGGTGCCGAGGCC 5909
Db 5635 CTGCTGACGGGCGGCGACACTGTGAGCTGCTTGGCGCGCTGCGCGGTGCCGAGGCC 5694
QY 5910 CAGGTTGCCAGACCGCTGAGCTGAGCTGCGGCTGCTGGACTCTCATGTATGCGACAG 5969
Db 5695 CAGGTTGCCAGACCGCTGAGCTGAGCTGCGGCTGCTGGACTCTCATGTATGCGACAG 5754
QY 5970 CGGCTGACGAGCACCCTACAGCGGAGGGAACAAACGCAAGCTGGGACCGCCCTGCGCTG 6029
Db 5755 CGGCTGACGAGCACCCTACAGCGGAGGGAACAAACGCAAGCTGGGACCGCCCTGCGCTG 5814
QY 6030 GTTGGGACCCAGCGCTGTTTCTGGAACGAGCGGACCAAGGATGGAACCCAGCGCGG 6089
Db 5815 GTTGGGACCCAGCGCTGTTTCTGGAACGAGCGGACCAAGGATGGAACCCAGCGCGG 5874
QY 6090 CGGCGCTTCTTTGGAACGAGCTTTTGGCGCTGCTGGGAGGCGCTTCACTGATGCTC 6149
Db 5875 CGGCGCTTCTTTGGAACGAGCTTTTGGCGCTGCTGGGAGGCGCTTCACTGATGCTC 5934
QY 6150 ACCTCCCATAGCATGAGGAGGTGTGAAGCGCTCTGCTCGGCCCTAGCCATCATGTTGAT 6209

Db 5935 ACCTCCCATAGCATGAGGAGGTGTGAAGCGCTCTCTGCGCCTAGCCATCATGTTGAT 5994
QY 6210 GGGCGGTTCCGCTCCCTGGGAGCGCCGCAACATCTCAAGGCGAGATTCCGCGCGGTCAC 6269
Db 5995 GGGCGGTTCCGCTCCCTGGGAGCGCCGCAACATCTCAAGGCGAGATTCCGCGCGGTCAC 6054
QY 6270 AACTGACCCCTGCGGGTCCCGCCGCAAGGTCCAGCGCGGCGCTTCTGTCGCGCC 6329
Db 6055 AACTGACCCCTGCGGGTCCCGCCGCAAGGTCCAGCGCGGCGCTTCTGTCGCGCC 6114
QY 6330 GAGTTCCCTGGGTCGGAGCTGCGGAGGCAATGAGAGCGCGCTTCCAGTTCGG 6389
Db 6115 GAGTTCCCTGGGTCGGAGCTGCGGAGGCAATGAGAGTCCGCTTCCAGTTCGG 6174
QY 6390 CCGGAGGCGCTCGGCCCTGGCGCGCTTTTGGAGAGCTGGCGGTGCAAGCGCGCAG 6449
Db 6175 CCGGAGGCGCTCGGCCCTGGCGCGCTTTTGGAGAGCTGGCGGTGCAAGCGCGCAG 6234
QY 6450 CACGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTC 6509
Db 6235 CACGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTTCTGTACTTC 6294
QY 6510 TCCAAGACACGCGGAAAGGACGAGGACACCGAAGACAGAGGAGCGAGTGGAGTG 6569
Db 6295 TCCAAGACACGCGGAAAGGACGAGGACACCGAAGACAGAGGAGCGAGTGGAGTG 6354
QY 6570 GACCCCGCGCGAGCGCTGCGACACCCAAACGCTGACGAGTTCCTCGATGACCCTAGC 6629
Db 6355 GACCCCGCGCGAGCGCTGCGACACCCAAACGCTGACGAGTTCCTCGATGACCCTAGC 6414
QY 6630 ACTCCGAGACTGTGCTC 6647
Db 6415 ACTCCGAGACTGTGCTC 6432

RESULT 12

US-10-775-920-6
; Sequence 6, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 6174
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-6

Query Match 89.3%; Score 6073.6; DB 17; Length 6174;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6082; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 702 CGCAGCGAATCCCTGGGGTTGGCACTGGCCCAAGCCAGGAGCCCTTGCAAGCTTTG 761
Db 79 CCCAGGAATCCCTGGGGTTGGCACTGGCCCAAGCCAGGAGCCCTTGCAAGCTTTG 138
QY 762 GAGCGCGCTGAGGACCTGGCCAGGAGCTCTGTCGCTGCGCAGCCTGGTGGAGTTCCG 821
Db 139 GAGCGCGCTGAGGACCTGGCCAGGAGCTCTGTCGCTGCGCAGCCTGGTGGAGTTCCG 198
QY 822 GCACTGTGACAGACCCCGAGGACCGAGCGGCCCTTGGAGTTGCTGTACAGGCCCTC 881
Db 199 GCACTGTGACAGACCCCGAGGACCGAGCGGCCCTTGGAGTTGCTGTACAGGCCCTC 258

QY	882	TGCAAGTGTACAGGGA	CCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACGAGGCTAGT	941
Db	259	TGCAAGTGTACAGGGA	CCTAGCAGCACAGTGGGCCCTCCCTCAACTGGTACGAGGCTAGT	318
QY	942	GACCTGATGAGCTGTGTGGGCA	AGCAGCAATTCGCGCTTCCGAGACAGCAGCCTGAGC	1001
Db	319	GACCTGATGAGCTGTGTGGGCA	AGCAGCAATTCGCGCTTCCGAGACAGCAGCCTGAGC	378
QY	1002	CCGCGCTGTCTCGAGCTGAT	TGAGAGCCCTGGAAGCAACCCGCTGTCCCGCTGTCTGTGG	1061
Db	379	CCGCGCTGTCTCGAGCTGAT	TGAGAGCCCTGGAAGCAACCCGCTGTCCCGCTGTCTGTGG	438
QY	1062	AGACGCTGAAGCTCTGAT	TCCTCGGGAAGCTACTTTTCACACAGATACACCTTTTACC	1121
Db	439	AGACGCTGAAGCTCTGAT	TCCTCGGGAAGCTACTTTTCACACAGATACACCTTTTACC	498
QY	1122	CGAAAGCTCTATGCGCCCA	AGGTGAACCGGACCTTCGAGAGCTCACCTCTGTGAGGGATGTC	1181
Db	499	CGAAAGCTCTATGCGCCCA	AGGTGAACCGGACCTTCGAGAGCTCACCTCTGTGAGGGATGTC	558
QY	1182	CGGAGGTGTGGAGATGCT	TGGAGACCCCGGATCTTACCTTATGAACGACAGTTCGAAT	1241
Db	559	CGGAGGTGTGGAGATGCT	TGGAGACCCCGGATCTTACCTTATGAACGACAGTTCGAAT	618
QY	1242	GTGSCCATGCTGACGCGCT	CTTCGAGATGACAGATGAGGAAGAGGAGCCCGACCT	1301
Db	619	GTGSCCATGCTGACGCGCT	CTTCGAGATGACAGATGAGGAAGAGGAGCCCGACCT	678
QY	1302	GGAGGCGGGACACAT	TGGAGGCGCTCGCATCTTTCTGACACCTTGGAGCGGTGGCTAC	1361
Db	679	GGAGGCGGGACACAT	TGGAGGCGCTCGCATCTTTCTGACACCTTGGAGCGGTGGCTAC	738
QY	1362	AGCTGGCAGACGACAC	CGCTGATGTGGGCACTGTGTGGGCAACGCTGGGCGAGTGACG	1421
Db	739	AGCTGGCAGACGACAC	CGCTGATGTGGGCACTGTGTGGGCAACGCTGGGCGAGTGACG	798
QY	1422	GAGTGCCTGCTTGGAC	AGCTGGAGCGGCAACCTTCAGAGCAGCCCTGTGTGTGCGG	1481
Db	799	GAGTGCCTGCTTGGAC	AGCTGGAGCGGCAACCTTCAGAGCAGCCCTGTGTGTGCGG	858
QY	1482	GCCTTGCAACTGTCT	CGCGGAACATCGATTTCTGGCGCGCGCTCGTCTTTCTGGACCTGAG	1541
Db	859	GCCTTGCAACTGTCT	CGCGGAACATCGATTTCTGGCGCGCGCTCGTCTTTCTGGACCTGAG	918
QY	1542	GACTCTTCAGACCCCA	CAGAGCACCCAACTGGGCGGCGGCGGCGGCGGCGGCGGCGG	1601
Db	919	GACTCTTCAGACCCCA	CAGAGCACCCAACTGGGCGGCGGCGGCGGCGGCGGCGGCGG	978
QY	1602	AAATCCGCATGACATT	GATGAGTGTGACGAGCAACCAATGAGATCAGGACAGGTTTGG	1661
Db	979	AAATCCGCATGACATT	GATGAGTGTGACGAGCAACCAATGAGATCAGGACAGGTTTGG	1038
QY	1662	GACCTTGGCCGAGCG	CGGACCCCTGACCGCATCTGGCTTACGCTGTGGGCGGCTTCTGTG	1721
Db	1039	GACCTTGGCCGAGCG	CGGACCCCTGACCGCATCTGGCTTACGCTGTGGGCGGCTTCTGTG	1098
QY	1722	TACTTGCAAGCTGT	TGGAGCGGTCCGCGTCTCAGCGGCGGCGGCGGCGGCGGCGG	1781
Db	1099	TACTTGCAAGCTGT	TGGAGCGGTCCGCGTCTCAGCGGCGGCGGCGGCGGCGGCGG	1158
QY	1782	GCCGCGCTCTACCT	GCAGGAGTGCCTATCCGCTGTATGTGACGACGTTCTTGGCT	1841
Db	1159	GCCGCGCTCTACCT	GCAGGAGTGCCTATCCGCTGTATGTGACGACGTTCTTGGCT	1218
QY	1842	GTGCTGAGCGCGT	CTGCTTCTGAGCTGCGCTTCTGAGCTGCGCTTCTGAGCTGAGCTG	1901
Db	1219	GTGCTGAGCGCGT	CTGCTTCTGAGCTGCGCTTCTGAGCTGCGCTTCTGAGCTGAGCTG	1278
QY	1902	ACAGTGAAGCGCT	GTGTGGGAGAGGAGACCGCGCTGCGGACCACTTCTGCGCGCATG	1961
Db	1279	ACAGTGAAGCGCT	GTGTGGGAGAGGAGACCGCGCTGCGGACCACTTCTGCGCGCATG	1338

QY	1962	GGCTCAGCGCGCGGT	GTCTGTAGGCTGGTTCCTCAGCTGCTCCTCGGCGCTTCTCTG	2021
Db	1339	GGCTCAGCGCGCGGT	GTCTGTAGGCTGGTTCCTCAGCTGCTCCTCGGCGCTTCTCTG	1398
QY	2022	CTCAGCGCGCGGT	GTCTGTAGGCTGGTTCCTCAGCTGGGGAATCTCTTCCCTTACAGCAC	2081
Db	1399	CTCAGCGCGCGGT	GTCTGTAGGCTGGTTCCTCAGCTGGGGAATCTCTTCCCTTACAGCAC	1458
QY	2082	CGGGCGTGGTCTT	CTTCTGTGTGGCAGGCTTTCGCGGTGGCCACGCTGACCCAGAGCTTC	2141
Db	1459	CGGGCGTGGTCTT	CTTCTGTGTGGCAGGCTTTCGCGGTGGCCACGCTGACCCAGAGCTTC	1518
QY	2142	CTGCTCAGCGCGCT	TTCTCTCCGCGCAACCTGGCTTGGCGCTTGGCGCGCTTGGCGCTAC	2201
Db	1519	CTGCTCAGCGCGCT	TTCTCTCCGCGCAACCTGGCTTGGCGCTTGGCGCGCTTGGCGCTAC	1578
QY	2202	TTCTCCTCTTAC	CTGCTGTGTGTGGCTTGTGGGACCGGCTTGGCGCGCTTGGCGGCT	2261
Db	1579	TTCTCCTCTTAC	CTGCTGTGTGTGGCTTGTGGGACCGGCTTGGCGCGCTTGGCGGCT	1638
QY	2262	GGCGCGTGGCGCG	AGCTGTCTCGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTG	2321
Db	1639	GGCGCGTGGCGCG	AGCTGTCTCGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTG	1698
QY	2322	GCTCTGCTGAGAG	CAGCGCGGCGGCTGAGTGCAAACTGGGCGCACCTGGGCTTACG	2381
Db	1699	GCTCTGCTGAGAG	CAGCGCGGCGGCTGAGTGCAAACTGGGCGCACCTGGGCTTACG	1758
QY	2382	GACAGCGTCTTCA	CGCTGTGCTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTT	2441
Db	1759	GACAGCGTCTTCA	CGCTGTGCTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTT	1818
QY	2442	GGCTTGGCGAC	CTGTGCTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2501
Db	1819	GGCTTGGCGAC	CTGTGCTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	1878
QY	2502	TGAAATTTTCT	TTCTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2561
Db	1879	TGAAATTTTCT	TTCTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	1938
QY	2562	CTTTGCCCAAC	CCCGCTTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2621
Db	1939	CTTTGCCCAAC	CCCGCTTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	1998
QY	2622	CTTGGCGTAT	TCCGCTTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2681
Db	1999	CTTGGCGTAT	TCCGCTTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2058
QY	2682	CGGGGCTCAG	CTGTGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2741
Db	2059	CGGGGCTCAG	CTGTGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2118
QY	2742	GCGGCAAGAC	CAACCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2801
Db	2119	GCGGCAAGAC	CAACCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2178
QY	2802	GCTTCTAT	TCCGCTTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2861
Db	2179	GCTTCTAT	TCCGCTTGGAGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGT	2238
QY	2862	GTCTGTCT	CAGTACAACTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2921
Db	2239	GTCTGTCT	CAGTACAACTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2298
QY	2922	TATGGCGGCT	GAAAGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2981
Db	2299	TATGGCGGCT	GAAAGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2358
QY	2982	CAGGATGTGGGCT	GTCTTCCAGAGAGAGTGTGGCTTGTGGCTTGTGGCTTGTGGCT	3041
Db	2359	CAGGATGTGGGCT	GTCTTCCAGAGAGAGTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2418
QY	3042	CAACGGAAGCT	GTCTTGGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	3101

Db 2419 CACCGAAGCTGTCTGGGCCAATTGGCTTTGTGGGGGCTCCCAAGTTGTTATCTGGAC 2478
QY 3102 GAGCCTACGGCTGGCGTGGATCTGTCTCCCGCGCGGTATTTGGGAGCTGTGTCTCAAA 3161
Db 2479 GAGCCTACGGCTGGCGTGGATCTGTCTCCCGCGCGGTATTTGGGAGCTGTGTCTCAA 2538
QY 3162 TACCGAAGAGTGGCAGCGTGAATCTCTCCACCAACACCTGGATGAGGAGAGTGTGTG 3221
Db 2539 TACCGAAGAGTGGCAGCGTGAATCTCTCCACCAACACCTGGATGAGGAGAGTGTGTG 2598
QY 3222 GAGACCGGTGGCTGTGGTGGAGTGGCGCGCTTGTGTGTGTGGCTCCCACTCTTC 3281
Db 2599 GAGACCGGTGGCTGTGGTGGAGTGGCGCGCTTGTGTGTGTGGCTCCCACTCTTC 2658
QY 3282 CTGCGCGCTCACTGGGCTCCGGCTACTACCTGAGCGTGGTGAAGCGCGCGCTGCGCCCTG 3341
Db 2659 CTGCGCGCTCACTGGGCTCCGGCTACTACTGACCTGAGTGAAGCGCGCGCTGCGCCCTG 2718
QY 3342 ACCACCAATGAGAGGCTGACACTGACATGAGAGGCAAGTGTGGACACAGGAGGAAAAG 3401
Db 2719 ACCACCAATGAGAGGCTGACACTGACATGAGAGGCAAGTGTGGACACAGGAGGAAAAG 2778
QY 3402 AAGATGGCAGCGGAGGAGAGTGGCAGCTCTCAGCTGCTGGCCCTGGTACAGCAC 3461
Db 2779 AAGATGGCAGCGGAGGAGGAGTGGCAGCTCTCAGCTGCTGGCCCTGGTACAGCAC 2838
QY 3462 TGGGTGCCCGGGGACGGCTGGTGGAGAGCTGGCCACAGAGCTGGTGGTGTGCGCC 3521
Db 2839 TGGGTGCCCGGGGACGGCTGGTGGAGAGCTGGCCACAGAGCTGGTGGTGTGCGCC 2898
QY 3522 TACACGGGTGCCCATGACGGCAGCTTCGCCACACTTTCGAGAGCTAGACACGGCGTG 3581
Db 2899 TACACGGGTGCCCATGACGGCAGCTTCGCCACACTTTCGAGAGCTAGACACGGCGTG 2958
QY 3582 GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAGGAGATCTTCCTG 3641
Db 2959 GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCCTCGAGGAGATCTTCCTG 3018
QY 3642 AAGTGTGGAGGAGTGTGTCGGACACAGATATGAGGAGTGGCAGCTGGGGGAGCAC 3701
Db 3019 AAGTGTGGAGGAGTGTGTCGGACACAGATATGAGGAGTGGCAGCTGGGGGAGCAC 3078
QY 3702 CTATGCACAGGATTTGGCTGAGACGTAAACCTTGGCGCTCAAGATGCCGCCACAGGAG 3761
Db 3079 CTATGCACAGGATTTGGCTGAGACGTAAACCTTGGCGCTCAAGATGCCGCCACAGGAG 3138
QY 3762 ACAGCGCTGGAGAACGGGGAAACAGTGGGTAGCCCCAGAGACTGACAGGGCTCTGGG 3821
Db 3139 ACAGCGCTGGAGAACGGGGAAACAGTGGGTAGCCCCAGAGACTGACAGGGCTCTGGG 3198
QY 3822 CCAGACCGGTGGGCGGGTACAGGGCTGGGGCACTGACCCGCCAGCAGCTCCAGGCCCTG 3881
Db 3199 CCAGACCGGTGGGCGGGTACAGGGCTGGGGCACTGACCCGCCAGCAGCTCCAGGCCCTG 3258
QY 3882 CTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGCGCTTGTTCGCCCAAGATCGTG 3941
Db 3259 CTCTCAAGCGCTTTCTGCTTGGCGCGCAGCGCGCGCGCTTGTTCGCCCAAGATCGTG 3318
QY 3942 CTGCTGCCCTTTTGTGGGCTGGGCCCTCTGTGTTTTCAGCTCATCTGTCCTCTTTTGGG 4001
Db 3319 CTGCTGCCCTTTTGTGGGCTGGGCCCTCTGTGTTTTCAGCTCATCTGTCCTCTTTTGGG 3378
QY 4002 CACTACCGGCTCTGCGGCTCAGTCCCAACATGATAGGCTGCTCAGGCTGCTTCTTCAGT 4061
Db 3379 CACTACCGGCTCTGCGGCTCAGTCCCAACATGATAGGCTGCTCAGGCTGCTTCTTCAGT 3438
QY 4062 GAGACGCCCGAGGGACCTCGAGCTGCCCGGCTGCTCGAGGCGCTGCTCGAGGGCA 4121
Db 3439 GAGACGCCCGAGGGACCTCGAGCTGCCCGGCTGCTCGAGGCGCTGCTCGAGGGCA 3498
QY 4122 GCACTGGAGGAGCCCCCAGTGCAGCATAGCTCCACAGGTTCTCGGACACAGAGTTCTCT 4181

Db 3499 GGACTGGAGGAGCCCCCAGTGCAGCATAGCTCCCAAGGTTCTCGGCACACAGAGTTCTCT 3558
QY 4182 GCTGAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCCAAGAGTCTCCATCCCCAGCC 4241
Db 3559 GCTGAAGTGGCCAAAGTCTTTGGCCAGTGGCAACTGGACCCCAAGAGTCTCCATCCCCAGCC 3618
QY 4242 TGCCAGTGTAGCCAGCCCGGTGCCCGCGCTGTGTCGCGAGCTGCCCGCTGCACTGCTGT 4301
Db 3619 TGCCAGTGTAGCCAGCCCGGTGCCCGCGCTGTGTCGCGAGCTGCCCGCTGCACTGCTGT 3678
QY 4302 GGTCCCTCCCTCGCCCGCAGGAGTGAACCGGCTCTGGGGAGTGGTTTCAGAACCTGACAGCC 4361
Db 3679 GGTCCCTCCCTCGCCCGCAGGAGTGAACCGGCTCTGGGGAGTGGTTTCAGAACCTGACAGCC 3738
QY 4362 CGGAACCTGTCTGACTTCTCTGCTCAAGACCTACCGCGCTGGTGGCGCAGGGCTGAGG 4421
Db 3739 CGGAACCTGTCTGACTTCTCTGCTCAAGACCTACCGCGCTGGTGGCGCAGGGCTGAGG 3798
QY 4422 ACTAAGAGTGGGTGAATGAGGTCAAGGTACGGAAGCTTCTCGTGGGGGGCGGAGACCCA 4481
Db 3799 ACTAAGAGTGGGTGAATGAGGTCAAGGTACGGAAGCTTCTCGTGGGGGGCGGAGACCCA 3858
QY 4482 GGCCTGCCCTCGGGCCCAAGAGTGGGGCGCTCAGTGGAGGAGTGTGGGGCGCTGCTGAGT 4541
Db 3859 GGCCTGCCCTCGGGCCCAAGAGTGGGGCGCTCAGTGGAGGAGTGTGGGGCGCTGCTGAGT 3918
QY 4542 CCCCTGCCCTCGGGGGCCCTCGACCGTCTGTAAGAAACCTCAGAGCTGGGCTCAGAGC 4601
Db 3919 CCCCTGCCCTCGGGGGCCCTCGACCGTCTGTAAGAAACCTCAGAGCTGGGCTCAGAGC 3978
QY 4602 CTGGATGCTCAGGAGCAGTCTCAAGATCTGGTTCAACAAAGGCTGGCACTCCATGGTG 4661
Db 3979 CTGGAGCTCAGGAGCAGTCTCAAGATCTGGTTCAACAAAGGCTGGCACTCCATGGTG 4038
QY 4662 GCTTTGTGAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCCGCCAGGCGCGGCC 4721
Db 4039 GCTTTGTGAACCGAGCCAGCAACGCAATCTCGTGTCTACCTGCCCGCCAGGCGCGGCC 4098
QY 4722 CGCCAGCCACAGCATCACCACTCAACCAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 4781
Db 4099 CGCCAGCCACAGCATCACCACTCAACCAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 4158
QY 4782 TCTGAGGCTGACATGATGGCTCTCGTGTGAGCGTCTCGTGTGCTCATCTGTGTGTCTTT 4841
Db 4159 TTTGAGGCTGACATGATGGCTCTCGTGTGAGCGTCTCGTGTGCTCATCTGTGTGTCTTT 4218
QY 4842 GCCATGCTTTTGTCCCGGCGAGCTTCACTTGTCTCATTTGAGGAGCGAGTCAACCGA 4901
Db 4219 GCCATGCTTTTGTCCCGGCGAGCTTCACTTGTCTCATTTGAGGAGCGAGTCAACCGA 4278
QY 4902 GCCAAGCACCTGACAGCTCATGGGGGGCTGTGCCCGACCTCTACTGCTTGGCACTTT 4961
Db 4279 GCCAAGCACCTGACAGCTCATGGGGGGCTGTGCCCGACCTCTACTGCTTGGCACTTT 4338
QY 4962 CTCTGGGACATGTGTAACTACTTGTGCGAGCATGCGTGTGCTCATCTTCTTGTGCC 5021
Db 4339 CTCTGGGACATGTGTAACTACTTGTGCGAGCATGCGTGTGCTCATCTTCTTGTGCC 4398
QY 5022 TTCCAGAGAGGGCATATGTGGCCCTGCCAACCTGCTCTCTGCTGTGTGCTACTA 5081
Db 4399 TTCCAGAGAGGGCATATGTGGCCCTGCCAACCTGCTCTCTGCTGTGTGCTACTA 4458
QY 5082 CTGTATGGCTGTGATCAGACCGCTCATGTACCCAGCTCTCTTCTTCTCTCGTGGCC 5141
Db 4459 CTGTATGGCTGTGATCAGACCGCTCATGTACCCAGCTCTCTTCTTCTCTCGTGGCC 4518
QY 5142 AGCACGCTATGTGGTGTCTCACTGTGATAAACCTTTTATTTGGCATCAATGGAGCATG 5201
Db 4519 AGCACGCTATGTGGTGTCTCACTGTGATAAACCTTTTATTTGGCATCAATGGAGCATG 4578
QY 5202 GCCACCTTTGTGTGCTGCTTCTCTGATCAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5261
Db 4579 GCCACCTTTGTGTGCTGCTTCTCTGATCAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4638

QY 5262 AAACAGGCTCTTCTTATCTTCCCACTTCTGCTTGGCGGGGCTCATTTGACATGGTG 5321
DB 4639 AAACAGGCTCTTCTTATCTTCCCACTTCTGCTTGGCGGGGCTTATTTGACATGGTG 4698
QY 5322 CGGAACAGGCCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGCGAGTTCCAGTCAACC 5381
DB 4699 CGGAACAGGCCATGGCTGATGCTTTGAGCGCTTGGGAGACAGGCGAGTTCCAGTCAACC 4758
QY 5382 CTGGGCTGGAGGTGGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTTC 5441
DB 4759 CTGGGCTGGAGGTGGTGGCAAGAACCTCTTGGCCATGTGTATACAGGGGCCCTCTTC 4818
QY 5442 CTTCTCTTCACTACTGCTGACACCGAAGCCAACTCTGCTGCCACAGCCACAGGCTGAGG 5501
DB 4819 CTTCTCTTCACTACTGCTGACACCGAAGCCAACTCTGCTGCCACAGCCACAGGCTGAGG 4878
QY 5502 TCTTGCCACTCTCTGGAGAGGAGGACGAGGATGTAGCCCTGTGAACGGGAGCGGCTGGTC 5561
DB 4879 TCTTGCCACTCTCTGGAGAGGAGGACGAGGATGTAGCCCTGTGAACGGGAGCGGCTGGTC 4938
QY 5562 CAAGGACCAACCAAGGGGATGTGTGCTGAGGAACTTGACCAAGGATATACCGTGGG 5621
DB 4939 CAAGGACCAACCAAGGGGATGTGTGCTGAGGAACTTGACCAAGGATATACCGTGGG 4998
QY 5622 CAGAGGATGCAGCTGTGACCGCTGTGCTGGGATTCCTCTGCTGAGTGTGTTTGGG 5681
DB 4999 CAGAGGATGCAGCTGTGACCGCTGTGCTGGGATTCCTCTGCTGAGTGTGTTTGGG 5058
QY 5682 CTGCTGGGTGTGAATGAGCAGGGAAGACGTCACGCTTTGCGCATGTGTGACGGGGACACA 5741
DB 5059 CTGCTGGGTGTGAATGAGCAGGGAAGACGTCACGCTTTGCGCATGTGTGACGGGGACACA 5118
QY 5742 TTGSCCAGACGGGCGAGGCTGTGCTGGCAGGCGCACAGCGTGGCCCGGGAACCCAGTGTCT 5801
DB 5119 TTGSCCAGACGGGCGAGGCTGTGCTGGCAGGCGCACAGCGTGGCCCGGGAACCCAGTGTCT 5178
QY 5802 GCGCACCTCAGCATGGGATACTGCCCTCAATCGATGCCATCTTTGAGCTGCTGACGGGC 5861
DB 5179 GCGCACCTCAGCATGGGATACTGCCCTCAATCGATGCCATCTTTGAGCTGCTGACGGGC 5238
QY 5862 CGCGAGCACCTGGAGCTGTGTCGCGCTGTCGCGGTGTCCGGAGGCCACAGTTTGCCAG 5921
DB 5239 CGCGAGCACCTGGAGCTGTGTCGCGCTGTCGCGGTGTCCGGAGGCCACAGTTTGCCAG 5298
QY 5922 ACCGCTGGCTCAGGCTGGCGGCTGTGGGACTCTCATGTGTACGACAGACCGGCTGCAGGC 5981
DB 5299 ACCGCTGGCTCAGGCTGGCGGCTGTGGGACTCTCATGTGTACGACAGACCGGCTGCAGGC 5358
QY 5982 ACCTACAGCGGAGGAACCAACGCMAGCTGGCGACGCCCTGGGCTGTGGGACCCA 6041
DB 5359 ACCTACAGCGGAGGAACCAACGCMAGCTGGCGACGCCCTGGGCTGTGGGACCCA 5418
QY 6042 GCCGTGTGTTCTTGGACGAGCGGACACACAGGATGGACCCACGCGCGCGCTTCCCTT 6101
DB 5419 GCCGTGTGTTCTTGGACGAGCGGACACAGGATGGACCCACGCGCGCGCTTCCCTT 5478
QY 6102 TGGAAACAGCCTTTTGGCCGTGGTGGGAGGCGGCTTCACTGATGTCTCACTCCCATAGC 6161
DB 5479 TGGAAACAGCCTTTTGGCCGTGGTGGGAGGCGGCTTCACTGATGTCTCACTCCCATAGC 5538
QY 6162 ATGGAGAGTGTGAGCGCTCTGCTCGCGCTAGCCATCATGTGTATGGGCGGTTCCGC 6221
DB 5539 ATGGAGAGTGTGAGCGCTCTGCTCGCGCTAGCCATCATGTGTATGGGCGGTTCCGC 5598
QY 6222 TGCTTGGGACCGCGCAACATCTCAAGGCGAGATTCCGCGGGGTACACACTGACCCCTG 6281
DB 5599 TGCTTGGGACCGCGCAACATCTCAAGGCGAGATTCCGCGGGGTACACACTGACCCCTG 5658
QY 6282 CGGCTGCCCGCGCAAGGTCACAGCGGACCGGCTTCTGTGGCGGCGGAGTTCCCTGGG 6341
DB 5659 CGGCTGCCCGCGCAAGGTCACAGCGGACCGGCTTCTGTGGCGGCGGAGTTCCCTGGG 5718

QY 6342 TCGAGCTGCGGAGGCACATGAGGCCGCTCGCTTCCAGCTGCCCGGAGGCGC 6401
DB 5719 TCGAGCTGCGGAGGCACATGAGGCCGCTCGCTTCCAGCTGCCCGGAGGCGC 5778
QY 6402 TCGCCCTGCGCGCGCTCTTTTGGAGAGCTGGCGGTGCACGGCGCAGAGCAACGCGGTGGAG 6461
DB 5779 TCGCCCTGCGCGCGCTCTTTTGGAGAGCTGGCGGTGCACGGCGCAGAGCAACGCGGTGGAG 5838
QY 6462 GACTTTTCGTGAGCCAGACGATGCTGGAGGAGGTATCTTTGACTTCTCAAGGACCA 6521
DB 5839 GACTTTTCGTGAGCCAGACGATGCTGGAGGAGGTATCTTTGACTTCTCAAGGACCA 5898
QY 6522 GGAAGAGCAGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCCA 6581
DB 5899 GGAAGAGCAGAGGACACCGAAGAGCAGAGGAGGAGGAGTGGAGTGGACCCCGGCCA 5958
QY 6582 GGCTCGAGCACCCCAACCGCTCAGCCAGTTCCTCGATGACCCCTAGCACTGCGAGACT 6641
DB 5959 GGCTCGAGCACCCCAACCGCTCAGCCAGTTCCTCGATGACCCCTAGCACTGCGAGACT 6018
QY 6642 GTCCTCTGAGCCTCCCTCCCTCGCGGCGCGCGGGAGGCCCTGGGAATGGCAAGGCAA 6701
DB 6019 GTCCTCTGAGCCTCCCTCCCTCGCGGCGCGCGGGAGGCCCTGGGAATGGCAAGGCAA 6078
QY 6702 GGTAGAGTGCCTAGGAGCCTGAGACTCAGGCTGGCAGAGGGGCTGGTGCCTCGAGAAAA 6761
DB 6079 GGTAGAGTGCCTAGGAGCCTGAGACTCAGGCTGGCAGAGGGGCTGGTGCCTCGAGAAAA 6138
QY 6762 TAAAGAGAGGCTGGAGAGACCGCTGCTGGTGAAA 6797
DB 6139 TAAAGAGAGGCTGGAGAGAGCGCTGGTGAAA 6174

RESULT 13

US-10-114-270-175
; Sequence 175, Application US/10114270
; Publication No. US20040030110A1

GENERAL INFORMATION:

; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Radigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Saha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086

QY	1890	TCGGTGACACTGACAGTGAAGCGCGTGGTGGGAGAGAGACGCGGCTCGGGACACC	1949
DB	1681	TCGGTGACACTGACAGTGAAGCGCGTGGTGGGAGAGAGACGCGGCTCGGGACACC	1740
QY	1950	ATGCGGCGCATATGGGGCTCAGCCGCGGGTGTCTCTGGCTAGGCTGGTTCTCAGCTGCTC	2009
DB	1741	ATGCGGCGCATATGGGGCTCAGCCGCGGGTGTCTCTGGCTAGGCTGGTTCTCAGCTGCTC	1800
QY	2010	GGGCGCTTCTGCTCAGCGCGCGCTGCTGGTTCTGGTGTCAAGCTGGGGACATCCCTC	2069
DB	1801	GGGCGCTTCTGCTCAGCGCGCGCTGCTGGTTCTGGTGTCAAGCTGGGGACATCCCTC	1860
QY	2070	CCCTACAGCACCCCGGCGGTGGTCTTCTCTGCTTCTGGCAGCCCTCGCGGTGGCCACGGTG	2129
DB	1861	CCCTACAGCACCCCGGCGGTGGTCTTCTCTGCTTCTGGCAGCCCTCGCGGTGGCCACGGTG	1920
QY	2130	ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCCAACTGGCTGGGCTCGGCG	2189
DB	1921	ACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCCAACTGGCTGGGCTCGGCG	1980
QY	2190	GGCGTGGCTTACTTCTCCCTTACTGCTGCTTACTGCTGCTGCTGCTGGCGGACCGG	2249
DB	1981	GGCGTGGCTTACTTCTCCCTTACTGCTGCTTACTGCTGCTGCTGCTGGCGGACCGG	2040
QY	2250	CTGCCGCGGCTGCGCGCTGGCGGAGCTGCTGCTGCGCCGCTGGCTTTCGCTTCGCG	2309
DB	2041	CTGCCGCGGCTGCGCGCTGGCGGAGCTGCTGCTGCGCCGCTGGCTTTCGCTTCGCG	2100
QY	2310	TGGGAGAGCCTGCTCTGCTGAGGAGCAGGGCGAGGGCGCGCATGGCACAACTGGGCG	2369
DB	2101	TGGGAGAGCCTGCTCTGCTGAGGAGCAGGGCGAGGGCGCGCATGGCACAACTGGGCG	2160
QY	2370	ACCGGCGCTTACGGAGACGCTTTCAGCTGGCGCCAGTCTCTGGCCTTCTGCTGTGGAC	2429
DB	2161	ACCGGCGCTTACGGAGACGCTTTCAGCTGGCGCCAGTCTCTGGCCTTCTGCTGTGGAC	2220
QY	2430	GCGGCGCTTACGGCGCTGCGCACTGTACTGGAAGCTGTGTGCGCCAGTACGGG	2489
DB	2221	GCGGCGCTTACGGCGCTGCGCACTGTACTGGAAGCTGTGTGCGCCAGTACGGG	2280
QY	2490	ATCCCTGAACCATGGAAATTTCTTTTCGAGGAGCTTACTGCTGCGGACCTCGGCCCCC	2549
DB	2281	ATCCCTGAACCATGGAAATTTCTTTTCGAGGAGCTTACTGCTGCGGACCTCGGCCCCC	2340
QY	2550	AAGATGACAGCCCTTGCCTCCACCCCGCTGGACCCAAAGTGTCTGTAGAAAGGACCG	2609
DB	2341	AAGATGACAGCCCTTGCCTCCACCCCGCTGGACCCAAAGTGTCTGTAGAAAGGACCG	2400
QY	2610	CCCGGCTGAGTCTGCGGCTATCCGTTTCGAGCTGAGAGAGCGCTTTCCTGGAAGCCG	2669
DB	2401	CCCGGCTGAGTCTGCGGCTATCCGTTTCGAGCTGAGAGAGCGCTTTCCTGGAAGCCG	2460
QY	2670	CAGCAGCCCTTGGGGGGCTCAGCTTGAATTTCTTACAGGGCCACATCAGCGCTTCTCTG	2729
DB	2461	CAGCAGCCCTTGGGGGGCTCAGCTTGAATTTCTTACAGGGCCACATCAGCGCTTCTCTG	2520
QY	2730	GGCCACAAACCGGGCGGCAAGACACACCTGTTCATCTGTAGTGGCTCTTCCACCC	2789
DB	2521	GGCCACAAACCGGGCGGCAAGACACACCTGTTCATCTGTAGTGGCTCTTCCACCC	2580
QY	2790	AGTGTGGCTTGGCTTCACTTGGGCGCAGAGCTCGGCTCCAGCATGGCGCCCATCCGG	2849
DB	2581	AGTGTGGCTTGGCTTCACTTGGGCGCAGAGCTCGGCTCCAGCATGGCGCCCATCCGG	2640
QY	2850	CCCCACCTGGGCGCTGTCTCTCAGTACAACAGTGTGTTTTCATGTGACCTGTGAGCGAG	2909
DB	2641	CCCCACCTGGGCGCTGTCTCTCAGTACAACAGTGTGTTTTCATGTGACCTGTGAGCGAG	2700
QY	2910	CAGTCTGGTCTATGGGCGGCTGAAGGGTCTGAGTGGCGGCTGTAGTGGCGCCCGAGAG	2969
DB	2701	CAGTCTGGTCTATGGGCGGCTGAAGGGTCTGAGTGGCGGCTGTAGTGGCGCCCGAGAG	2760
QY	2970	GACCGTCTGTCAGGAGTGTGGGGCTGGTCTCCAAAGCAGAGTGTGCAGACTCGGCCACTC	3029

DB	2761	GACCGTCTGTCGAGGATGTGGGGCTGGTCTCCAAAGCAGAGTGTGCAGACTCGCCACTC	2820
QY	3030	TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTTGCTTTGTGGCGGCTCCCAAGTT	3089
DB	2821	TCTGTGGGATGCAACGGAAGCTGTCCGTGGCCATTTGCTTTGTGGCGGCTCCCAAGTT	2880
QY	3090	GTTATCTCTGAGAGAGCTACGGCTGCGCTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG	3149
DB	2881	GTTATCTCTGAGAGAGCTACGGCTGCGCTGGATCTCTGCTTCCCGCGCGGTATTTGGGAG	2940
QY	3150	CTGCTGCTCAAAATACCGAGAAGTGCACGCTGATCTCTCCACCCACCACTGGATGAG	3209
DB	2941	CTGCTGCTCAAAATACCGAGAAGTGCACGCTGATCTCTCCACCCACCACTGGATGAG	3000
QY	3210	GCAGAGCTCTGGGAGACCGTGTGGCTGTGGTGGAGGTGGCGCTGTGCTGCTGGC	3269
DB	3001	GCAGAGCTCTGGGAGACCGTGTGGCTGTGGTGGAGGTGGCGCTGTGCTGCTGGC	3060
QY	3270	TCCCACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC	3329
DB	3061	TCCCACTCTTCTGCGCGCTCACCTGGGCTCCGGCTACTACTGACGCTGGTGAAGGCC	3120
QY	3330	CGCTGCTCCCTGACCAACCAATGAGAAGTGCACCTGACATGAGGAGGAGTGTGGACAC	3389
DB	3121	CGCTGCTCCCTGACCAACCAATGAGAAGTGCACCTGACATGAGGAGGAGTGTGGACAC	3180
QY	3390	AGGCGGAAAAGAAAGTGGCAGCGGCGAGGAGTGGGCACTCTCAGCTGTGGCC	3449
DB	3181	AGGCGGAAAAGAAAGTGGCAGCGGCGAGGAGTGGGCACTCTCAGCTGTGGCC	3240
QY	3450	CTGGTACAGCTCTGGTGGCGGCGAGCTGGTGGAGGAGTGCACACAGAGCTGGTG	3509
DB	3241	CTGGTACAGCTCTGGTGGCGGCGAGCTGGTGGAGGAGTGCACACAGAGCTGGTG	3300
QY	3510	CTGGTCTCTCCCTTACACGGGTCCTATGACGGGAGCTTCCGCACTCTTCCGAGAGCTA	3569
DB	3301	CTGGTCTCTCCCTTACACGGGTCCTATGACGGGAGCTTCCGCACTCTTCCGAGAGCTA	3360
QY	3570	GACACGCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAG	3629
DB	3361	GACACGCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGGCTCGAG	3420
QY	3630	GAGATCTTCTGAAAGTGTGGAGGAGTGTCTGCGGACACAGATATGAGGAGTGCAGC	3689
DB	3421	GAGATCTTCTGAAAGTGTGGAGGAGTGTCTGCGGACACAGATATGAGGAGTGCAGC	3480
QY	3690	TGCGGCGAGCACCTATGACAGGCAATTTGCTGGCTTAGACGTAAACCTGCGGCTCAAGATG	3749
DB	3481	TGCGGCGAGCACCTATGACAGGCAATTTGCTGGCTTAGACGTAAACCTGCGGCTCAAGATG	3540
QY	3750	CGGCCACAGAGACAGCGGTGGAGAACGGGAAACAGAGTGGGTGAGCCCGCAGAGCTGAC	3809
DB	3541	CGGCCACAGAGACAGCGGTGGAGAACGGGAAACAGAGTGGGTGAGCCCGCAGAGCTGAC	3600
QY	3810	CAGGCTCTGGGCGCAGACCGCTGGGCGGCTTACAGGCTGGGCACTGACCCCGCAGAG	3869
DB	3601	CAGGCTCTGGGCGCAGACCGCTGGGCGGCTTACAGGCTGGGCACTGACCCCGCAGAG	3660
QY	3870	CTCCAGGCGCTGCTTCTCAAGCGCTTCTGCTTTCGCGCGCAGCGCGCGGCTGTTC	3929
DB	3661	CTCCAGGCGCTGCTTCTCAAGCGCTTCTGCTTTCGCGCGCAGCGCGGCTGTTC	3720
QY	3930	GCCAGAGCTGCTGCTGCTCTTTTGGGCGCTGGCGCTGGTTCAGCTCATCTG	3989
DB	3721	GCCAGAGCTGCTGCTGCTCTTTTGGGCGCTGGCGCTGGTTCAGCTCATCTG	3780
QY	3990	CTTCTTTTGGGCGCTTACCGGCTCTGGGCTCAGTCCCGCAGCATGACCTGCTCAGGTG	4049
DB	3781	CTTCTTTTGGGCGCTTACCGGCTCTGGGCTCAGTCCCGCAGCATGACCTGCTCAGGTG	3840
QY	4050	TCTTCTTTCAGTGAAGACGCGCCCGAGGGGACCTTGGAGCTGCGCGGCTGCTGAGGGCGCTG	4109

Qy	6270	ACACTGACCCTTCGGGGTCCCGCGCAAGGTCCAGCCGGCAGCGGCCTTCGTGGCGGCC	6329
Db	5800	ACA CTGACCCTTCGGGGTCCCGCGCAAGGTCCAGCCGGCAGCGGCCTTCGTGGCGGCC	5859
Qy	6330	GAGTTCCCTGGGTCCGAGCTCCGCGAGGCAATGGAGGCCCTTGCGCTTCCAGCTGCCG	6389
Db	5860	GAGTTCCCTGGGTCCGAGCTCCGCGAGGCAATGGAGGCCCTTGCGCTTCCAGCTGCCG	5919
Qy	6390	CCGGAGGGCGCTGCGCCCTGGCGCGCTCTTTGGAGAGCTGGCGGTGCACGCGCCAGAG	6449
Db	5920	CCGGAGGGCGCTGCGCCCTGGCGCGCTCTTTGGAGAGCTGGCGGTGCACGCGCCAGAG	5979
Qy	6450	CACGGCGTGGAGCATTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTCTTTGTACTTC	6509
Db	5980	CACGGCGTGGAGCATTTTCCGTGAGCCAGACGATGCTGGAGGAGTATTCTTTGTACTTC	6039
Qy	6510	TCCAGAGCACAGGGGAAGGACGAGGACACCGAAGAGCAGAAGGAGGAGGATGGAGTG	6569
Db	6040	TCCAAGGACAGGGGAAGGACGAGGACACCGAAGAGCAGAAGGAGGAGGATGGAGTG	6099
Qy	6570	GACCCCGCGCAGGCGCTGCAGCACCCCAAACGGGTACGCCAGTTCCTCCGATGACCCTAGC	6629
Db	6100	GACCCCGCGCAGGCGCTGCAGCACCCCAAACGGGTACGCCAGTTCCTCCGATGACCCTAGC	6159
Qy	6630	ACTGCCGAGACTGTGCTCTGAGCCTCCCTCCCTCGCGGGCCGCGGGAGGCCCTCGGAA	6689
Db	6160	ACTGCCGAGACTGTGCTCTGAGCCTCCCTCCCTCGCGGGCCGCGGGAGGCCCTCGGAA	6219
Qy	6690	TGGCAAGGGCAAGGTAGAGTGCTTAGAGGCCCTTGGA CTCAGGCTGGCAGAGGGCTGGTG	6749
Db	6220	TGGCAAGGGCAAGGTAGAGTGCTTAGAGGCCCTTGGA CTCAGGCTGGCAGAGGGCTGGTG	6279
Qy	6750	CCCTGGAGAAAATAAAGAGAGGCTGGAGAGAGCCGCTGCTGGTGAA	6797
Db	6280	CCCTGGAGAAAATAAAGAGAGGCTGGAGAGAGCCGCTGCTGGTGAA	6327

RESULT 14
US-10-775-920-5
; Sequence 5, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; TITLE OF INVENTION: IN CERTAIN CANCERS
; FILE REFERENCE: Mergen - 00108
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 6027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-5

Query Match		87.1%;	Score 5928.2;	DB 17;	Length 6027;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 5936;		Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;

Qy	702	CGACCGAAATCCCTGGGGTTGGCACTGGGGCAAGCCAGAGCCCTTGCA CAGCTTGTG	761
Db	79	CCCACGAAATCCCTGGGGTTGGCACTGGGGCAAGCCAGAGCCCTTGCA CAGCTTGTG	138
Qy	762	GAGCGCGCTCAGGACCTGGCCCCAGGAGCTCCTGGCGCTGCGAGCTGGTGGAGCTCGG	821
Db	139	GAGCGCGCTCAGGACCTGGCCCCAGGAGCTCCTGGCGCTGCGAGCTGGTGGAGCTTCGG	198
Qy	822	GCACTGCTCAGACAGACCCCGAGGGAGCACAGCGGGCCCCCTGGAGTTGCTGT CAGAGGCCCTC	881

1962 QY GGGCTCAGCCGCGCGGTGCTCTGGCTAGGCTGTTCTCTCAGCTGCTCGGGCCTTCCTG 2021
1339 Db GGGCTCAGCCGCGCGGTGCTCTGGCTAGGCTGTTCTCTCAGCTGCTCGGGCCTTCCTG 1398
2022 QY CTCAGGCGCGCTGCTGGTCTGCTCAAGCTGGGGACATCTCTCCCTACAGCCAC 2081
1399 Db CTCAGGCGCGCTGCTGGTCTGCTCAAGCTGGGAGACATCTCTCCCTACAGCCAC 1458
2082 QY CCGGCGGTGCTCTCTGCTTCTGGCAGCCTTCGCGGTGGCCACCGGTGACCCAGAGCTTC 2141
1459 Db CCGGCGGTGCTCTCTGCTTCTGGCAGCCTTCGCGGTGGCCACCGGTGACCCAGAGCTTC 1518
2142 QY CTGCTCAGCGCTTCTTCTCCGCGCAACCTGGCTGGCGCTCGGGCGCTGCGCTAC 2201
1519 Db CTGCTCAGCGCTTCTTCTCCGCGCAACCTGGCTGGCGCTCGGGCGCTGCGCTAC 1578
2202 QY TTCTCCCTCTACCTGCGCTTACGCTGCTGTGTGGCTTGGCGGACCGGCTGCCGCGGT 2261
1579 Db TTCTCCCTCTACCTGCGCTTACGCTGCTGTGTGGCTTGGCGGACCGGCTGCCGCGGT 1638
2262 QY GCGCGCTGGCGGCGAGCTGCTGCTGCGCGGTGCTTCCGCTTCGGCTGGAGAGCCTG 2321
1639 Db GCGCGCTGGCGGCGAGCTGCTGCTGCGCGGTGCTTCCGCTTCGGCTTCGGCTTCGGAGCGCTG 1698
2322 QY GCTCTGCTGGAGGACGAGCGGCGGCGCAAGTGGGCAACGCTGGGCAACCGGCTTACG 2381
1699 Db GCTCTGCTGGAGGACGAGCGGCGGCGCAAGTGGGCAACGCTGGGCAACCGGCTTACG 1758
2382 QY CGAGAGCTTTCAGCTGCGCCAGCTCTGCGCTTCTGCTGGAGCGCGGCTCTAC 2441
1759 Db CGAGAGCTTTCAGCTGCGCCAGCTCTGCGCTTCTGCTGGAGCGCGGCTCTAC 1818
2442 QY GSCCTCGCACCTGCTACTGCGAAGCTGTGTGCCAGGCGAGTACGCGGATCCCTGAACCA 2501
1819 Db GSCCTCGCACCTGCTACTGCGAAGCTGTGTGCCAGGCGAGTACGCGGATCCCTGAACCA 1878
2502 QY TGGAAATTTTCTTCGAGGAGCTACTGCTGCGGACCTCGGCCCCCAAGAGTCCAGCC 2561
1879 Db TGGAAATTTTCTTCGAGGAGCTACTGCTGCGGACCTCGGCCCCCAAGAGTCCAGCC 1938
2562 QY CTTGCGCCACCCGCTGGAACCAAGTGTGTGTAAGAGGCAACCGCGGCTGAGT 2621
1939 Db CTTGCGCCACCCGCTGGAACCAAGTGTGTGTAAGAGGCAACCGCGGCTGAGT 1998
2622 QY CTTGCGCTATCCGTTGCGAGCTGGAAGCGCTTCTGGAAGCGCCGAGCGCCCTG 2681
1999 Db CTTGCGCTATCCGTTGCGAGCTGGAAGCGCTTCTGGAAGCGCCGAGCGCCCTG 2058
2682 QY CCGGCGCTCAGCTGGAATTTCTACAGGCGCCACATCACCGCTTCTCGGCGCAACAGCGG 2741
2059 Db CCGGCGCTCAGCTGGAATTTCTACAGGCGCCACATCACCGCTTCTCGGCGCAACAGCGG 2118
2742 QY GCGGCAAGACCAACCCCTGTCATCTGAGTGGCTCTTCCACCAAGTGGTCT 2801
2119 Db GCGGCAAGACCAACCCCTGTCATCTGAGTGGCTCTTCCACCAAGTGGTCT 2178
2802 QY GCCTTCATCTGGGCGACGAGCTCGCTCCAGCATGGCGGCATCCGSCCCCACTGGGC 2861
2179 Db GCCTTCATCTGGGCGACGAGCTCGCTCCAGCATGGCGGCATCCGSCCCCACTGGGC 2238
2862 QY GTCTGCTCAGTACAAAGCTGTGTTGACATGCTGACCGTGGAGAGCAGTCTGGTTC 2921
2239 Db GTCTGCTCAGTACAAAGCTGTGTTGACATGCTGACCGTGGAGAGCAGTCTGGTTC 2298
2922 QY TATGGCGGCTGAAGGCTCTGAGTCCGCTGTAGTGGCCCCCGAGAGGACCTGTGCTG 2981
2299 Db TATGGCGGCTGAAGGCTCTGAGTCCGCTGTAGTGGCCCCCGAGAGGACCTGTGCTG 2358
2982 QY CAGGATGTGGGCTGGTCTCCAGCAGAGTGTGACAGTCCGCACTCTCTGTGGGATG 3041
2359 Db CAGGATGTGGGCTGGTCTCCAGCAGAGTGTGACAGTCCGCACTCTCTGTGGGATG 2418

3042 QY CAACGGAAGCTGCTCGTGGCCATTGCTTTGTGGCGGCTCCCAAGTTGTTATCTCGGAC 3101
2419 Db CAACGGAAGCTGCTCGTGGCCATTGCTTTGTGGCGGCTCCCAAGTTGTTATCTCGGAC 2478
3102 QY GAGCTACGGCTGGGCTGGATCCTGCTTCCCGCGGGTATTTGGAGCTGCTGTCAA 3161
2479 Db GAGCTACGGCTGGGCTGGATCCTGCTTCCCGCGGGTATTTGGAGCTGCTGTCAA 2538
3162 QY TACCGAAGCTCGCACGCTGATCTCTCCACCCACCACTGGATGAGCAGAGCTGCTG 3221
2539 Db TACCGAAGCTCGCACGCTGATCTCTCCACCCACCACTGGATGAGCAGAGCTGCTG 2598
3222 QY GGAGACCTGTGTGCTGGTGGCAGGTGGCGCTTGTGCTGTGTGCTGCCACTCTTC 3281
2599 Db GGAGACCTGTGTGCTGGTGGCAGGTGGCGCTTGTGCTGTGTGCTGCCACTCTTC 2658
3282 QY CTGGCGCTGACCTGGGCTCGGCTACTACTGAGCTGTGTGAAGGCGCGGCTGCCCTG 3341
2659 Db CTGGCGCTGACCTGGGCTCGGCTACTACTGAGCTGTGTGAAGGCGCGGCTGCCCTG 2718
3342 QY ACCACCAATGAGAAGGCTGACACTGACATGAGGAGCAGTGTGGACACAGGCGAGAAAG 3401
2719 Db ACCACCAATGAGAAGGCTGACACTGACATGAGGAGCAGTGTGGACACAGGCGAGAAAG 2778
3402 QY AAGAAATGCAAGGCGGAGCAGAGTGGCACTCTCTAGCTGTGTGGCGCTTGTACAGCAC 3461
2779 Db AAGAAATGCAAGGCGGAGCAGAGTGGCACTCTCTAGCTGTGTGGCGCTTGTACAGCAC 2838
3462 QY TGGGTGCCCGGGGACCGGCTGGTGGAGGCTGCCACAGAGCTGTGTGTGTGCC 3521
2839 Db TGGGTGCCCGGGGACCGGCTGGTGGAGGCTGCCACAGAGCTGTGTGTGTGCC 2898
3522 QY TACACGGGTGCCATGACGGCAGCTTCCGCCACACTCTTCCGAGAGCTAGACACCGGGCTG 3581
2899 Db TACACGGGTGCCATGACGGCAGCTTCCGCCACACTCTTCCGAGAGCTAGACACCGGGCTG 2958
3582 QY CCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCTCGAGGAGATCTTCTG 3641
2959 Db CCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACAGCTCGAGGAGATCTTCTG 3018
3642 QY AAGTGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGAGCTGGGGCAGCAC 3701
3019 Db AAGTGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGAGCTGGGGCAGCAC 3078
3702 QY CTATGCAAGGCTATGCTGCTGCTAGACTACCTCGGCTCAAGATGCCGCCACAGGAG 3761
3079 Db CTATGCAAGGCTATGCTGCTGCTAGACTACCTCGGCTCAAGATGCCGCCACAGGAG 3138
3762 QY ACAGCGCTGGAGAACGGGGAACAGCTGGGTCAGCCCCCAGAGCTGACAGGGCTCTGGG 3821
3139 Db ACAGCGCTGGAGAACGGGGAACAGCTGGGTCAGCCCCCAGAGCTGACAGGGCTCTGGG 3198
3822 QY CCAGACCGCTGGGCGCGGCTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTG 3881
3199 Db CCAGACCGCTGGGCGCGGCTACAGGGCTGGGCACTGACCCCGCAGAGCTCCAGGCGCTG 3258
3882 QY CTTCTCAAGCGCTTCTGCTTGGCGCGAGCGCGCGGCTGTCGCCAGATCTG 3941
3259 Db CTTCTCAAGCGCTTCTGCTTGGCGCGAGCGCGCGGCTGTCGCCAGATCTG 3318
3942 QY CTGCTGCGCTTCTTGTGGGCTGGCGCTCGTGTTCAGGCTTCATCGTCTCTTTCGGG 4001
3319 Db CTGCTGCGCTTCTTGTGGGCTGGCGCTCGTGTTCAGGCTTCATCGTCTCTTTCGGG 3378
4002 QY CACTACCGGCTCTGCGGCTCAGTCCACCATGATGACGGTGTCTAGGTGTCTTCTTCAGT 4061
3379 Db CACTACCGGCTCTGCGGCTCAGTCCACCATGATGACGGTGTCTAGGTGTCTTCTTCAGT 3438
4062 QY GAGGACCGCCAGGGGACCTTGGAGCTGCCCGGCTGCTCGAGGCGCTGCTCGAGGAGCA 4121
3439 Db GAGGACCGCCAGGGGACCTTGGAGCTGCCCGGCTGCTCGAGGCGCTGCTCGAGGAGCA 3498
4122 QY GGACTGGAGGAGCCCCCAGTGTCAGCATAGTCTCCACAGGTTCTCGGACACAGAAAGTTCCT 4181

Db 3499 GGAAGGAGGAGCCCGAGTGCAGCATAGCTCCACAGGTTCTCGGCACCAAGAGTTCTCT 3558
QY 4182 GCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGACCCCAAGAGTCTCCATCCCAAGCC 4241
Db 3559 GCTGAAGTGGCCAAAGTCTTGGCCAGTGGCACTGGACCCCAAGAGTCTCCATCCCAAGCC 3618
QY 4242 TGGCAGTGTAGCAGCCGCGTCCCGCGCTGCTGCTGCCCACTGCGCCGGCTGCAGCTGGT 4301
Db 3619 TGGCAGTGTAGCAGCCGCGTCCCGCGCTGCTGCTGCCCACTGCGCCGGCTGCAGCTGGT 3678
QY 4302 GGTCCCGCTCCCGCCCAAGGAGTGAACCGGCTCTGGGAAAGTGGTTCAGAACTCTGACAGGC 4361
Db 3679 GGTCCCGCTCCCGCCCAAGGAGTGAACCGGCTCTGGGAAAGTGGTTCAGAACTCTGACAGGC 3738
QY 4362 CGGAACTGTGTGATCTTCTGTGTCAAGACCTACCCGCGCTGTGTGCCCAAGGCGCTGAAG 4421
Db 3739 CGGAACTGTGTGATCTTCTGTGTCAAGACCTACCCGCGCTGTGTGCCCAAGGCGCTGAAG 3798
QY 4422 ACTAAGAGTGGGTGAATGAGTCAAGTACGAGGCTTCTGCTGGGGGCCGAGACCCA 4481
Db 3799 ACTAAGAGTGGGTGAATGAGTCAAGTACGAGGCTTCTGCTGGGGGCCGAGACCCA 3858
QY 4482 GGCCTGCGCTCGGCGCAAGAGTTGGCGCTCAAGTGGAGAGTGTGGCGCTGCTGAGT 4541
Db 3859 GGCCTGCGCTCGGCGCAAGAGTTGGCGCTCAAGTGGAGAGTGTGGCGCTGCTGAGT 3918
QY 4542 CCCTGCTGCGGGGCCCTGACCGTGTCTGTAAGAACTCAAGCTGGGCTCAAGC 4601
Db 3919 CCCTGCTGCGGGGCCCTGACCGTGTCTGTAAGAACTCAAGCTGGGCTCAAGC 3978
QY 4602 CTGATGCTCAGGACAGTCTCAGATCTGGTTCAAGAAAGCTGGCACTCCATGGTG 4661
Db 3979 CTGATGCTCAGGACAGTCTCAGATCTGGTTCAAGAAAGCTGGCACTCCATGGTG 4038
QY 4662 GCCTTTGTCAACCGAGCCAGCAAGCAATCTCGTGTCTCACTGCCCGCCAGCGCGGCC 4721
Db 4039 GCCTTTGTCAACCGAGCCAGCAAGCAATCTCGTGTCTCACTGCCCGCCAGCGCGGCC 4098
QY 4722 CGGCACGCCCAAGATCAACACCTCAACACCCCTTGAACTCAACAGGAGCAGCTG 4781
Db 4099 CGGCACGCCCAAGATCAACACCTCAACACCCCTTGAACTCAACAGGAGCAGCTG 4158
QY 4782 TCTGAGGCTGCACTGATGGGCTCTCGGTGGAGTCTCTGCTCCATCTGTGTGTCTTT 4841
Db 4159 TTTGAGGCTGCACTGATGGGCTCTCTCGGTGGAGTCTCTGCTCCATCTGTGTGTCTTT 4218
QY 4842 GCCATGTCTTTGTCCCGGCCAGCTTCACTCTTGTCTCAATTGAGAGCGAGTCAACCCGA 4901
Db 4219 GCCATGTCTTTGTCCCGGCCAGCTTCACTCTTGTCTCAATTGAGAGCGAGTCAACCCGA 4278
QY 4902 GCCAAGCACTGAGCTCATGGGGGCTGTGCCCACTCTCTACTGCGCTTGGCACTTT 4961
Db 4279 GCCAAGCACTGAGCTCATGGGGGCTGTGCCCACTCTCTACTGCGCTTGGCACTTT 4338
QY 4962 CTCTGGGACATGTGTACTTGTGGTGGCAGATGATCGTGTGCTCACTTTCTGGCC 5021
Db 4339 CTCTGGGACATGTGTAACTTGTGGTGGCAGATGATCGTGTGCTCACTTTCTGGCC 4398
QY 5022 TTCCAGCAGAGGGCATATGTGGGCCCTTGGCCAACTGCTGCTCTCTGCTGTGTGCTACTA 5081
Db 4399 TTCCAGCAGAGGGCATATGTGGGCCCTTGGCCAACTGCTGCTCTCTGCTGTGTGCTACTA 4458
QY 5082 CTGTAATGGCTGGTGCATCAACCGCTCATGTAACCAAGCTCTCTTCTTCTCCGTGCC 5141
Db 4459 CTGTAATGGCTGGTGCATCAACCGCTCATGTAACCAAGCTCTCTTCTTCTCCGTGCC 4518
QY 5142 AGCAGAGCTTATGGTGTCTCACTGATAACTCTTTATTTGGCATCATGGAAGCATG 5201
Db 4519 AGCAGAGCTTATGGTGTCTCACTGATAACTCTTTATTTGGCATCAATGGAAGCATG 4578
QY 5202 GCCACCTTTGTGTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTGAGCCGATCTTG 5261

Db 4579 GCCACCTTTGTGCTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTGAGCCGATCTTG 4638
QY 5262 AAACAGGTCTTCTTATCTTCCCGCACTTCTGCTTGGCGGGGGCTCATTTGACATGGTG 5321
Db 4639 AAACAGGTCTTCTTATCTTCCCGCACTTCTGCTTGGCGGGGGCTTATTTGACATGGTG 4698
QY 5322 CGGAAACAGGCCCATGGCTGATGCTTTTGTAGCGCTTGGGAGACAGGCAAGTTCCTAGTACCCC 5381
Db 4699 CGGAAACAGGCCCATGGCTGATGCTTTTGTAGCGCTTGGGAGACAGGCAAGTTCCTAGTACCCC 4758
QY 5382 CTGCGTGGGAGTGTTCGGCAAGAACCTTCTTGGCCATGATGATACAGGGGCCCTTCTTC 5441
Db 4759 CTGCGTGGGAGTGTTCGGCAAGAACCTTCTTGGCCATGATGATACAGGGGCCCTTCTTC 4818
QY 5442 CTTTCTTTTACACTACTGCTGACGACCCGAAGCACTCTCTGCCACAGCCCAAGGTTAGG 5501
Db 4819 CTTTCTTTTACACTACTGCTGACGACCCGAAGCACTCTCTGCCACAGCCCAAGGTTAGG 4878
QY 5502 TCTTCTGCCACTCTCGGAGAGGAGGACGAGGATGTAGCCCGTGAAACGGGAGCGGTTGGTC 5561
Db 4879 TCTTCTGCCACTCTCGGAGAGGAGGAGGAGGAGGATGTAGCCCGTGAAACGGGAGCGGTTGGTC 4938
QY 5562 CAAAGGAGCCACCCAGGGGAGTGTGTGGTGTCTGAGGAACCTTGACCAAGGTTATACCGTGGG 5621
Db 4939 CAAAGGAGCCACCCAGGGGAGTGTGTGGTGTCTGAGGAACTTGACCAAGGTTATACCGTGGG 4998
QY 5622 CAGAGGATGCCAGCTTGTGACCGCTTGTGCTTGGGATTTCCCGCTGGTGTGTTTGGG 5681
Db 4999 CAGAGGATGCCAGCTTGTGACCGCTTGTGCTTGGGATTTCCCGCTGGTGTGTTTGGG 5058
QY 5682 CTGCTGGGTGTGAATGTAGCAGGGAAGAAGCTTCCACGTTTTCGCAATGTGTAGCGGGGACACA 5741
Db 5059 CTGCTGGGTGTGAATGTAGCAGGGAAGAAGCTTCCACGTTTTCGCAATGTGTAGCGGGGACACA 5118
QY 5742 TTGGCCAGCAGGGGCGAGGCTGTGTGTGGCAGGACACAGCGTGGCCCGGGAACCCAGTGCT 5801
Db 5119 TTGGCCAGCAGGGGCGAGGCTGTGTGTGGCAGGACACAGCGTGGCCCGGGAACCCAGTGCT 5178
QY 5802 GGCACCTCAGCATGGGATCTGCGCTTCAATTCGGATGCCATCTTTGAGCTGTGACGGGC 5861
Db 5179 GGCACCTCAGCATGGGATCTGCGCTTCAATTCGGATGCCATCTTTGAGCTGTGACGGGC 5238
QY 5862 CGCGAGCACCTGAGGCTGTGTGCGCCCTGCGGGGTGTCCGAGGCCAGGTTGCCAG 5921
Db 5239 CGCGAGCACCTGAGGCTGTGTGCGCCCTGCGGGGTGTCCGAGGCCAGGTTGCCAG 5298
QY 5922 ACCGCTGGCTCAGGCTTGGCGGTCTTGGGACTCTCTATGGTACGACACCGGCTTGCAGGC 5981
Db 5299 ACCGCTGGCTCAGGCTTGGCGGTCTTGGGACTCTCTATGGTACGACACCGGCTTGCAGGC 5358
QY 5982 ACCTCAGCGGAGGGAACAAACGCAAGCTGGCGACCGGCCCTTGGCGCTGTGGGGACCCA 6041
Db 5359 ACCTCAGCGGAGGGAACAAACGCAAGCTGGCGACCGGCCCTTGGCGCTGTGGGGACCCA 5418
QY 6042 GCGTGGTGTCTTGACGAGCGGACACAGGATGGAACCCAGCGCGGGCTTCTCTT 6101
Db 5419 GCGTGGTGTCTTGACGAGCGGACACAGGATGGAACCCAGCGCGGGCTTCTCTT 5478
QY 6102 TGGAAACAGGCTTTTGGCCGCTGTGGGAGGGGCGTTTCAGTGTATGCTCACCTCCCAATAGC 6161
Db 5479 TGGAAACAGGCTTTTGGCCGCTGTGGGAGGGGCGTTTCAGTGTATGCTCACCTCCCAATAGC 5538
QY 6162 ATGGAGGATGTGAAGCGCTCTGCTCGCGCTAGGCCATCATGTTGTAATGGGCGGTTCCGC 6221
Db 5539 ATGGAGGATGTGAAGCGCTCTGCTCGCGCTAGGCCATCATGTTGTAATGGGCGGTTCCGC 5598
QY 6222 TGCTTGGCAGCCCGCAACATCTCAAGGCGAGATTCGCGGGGCTCACAACACTGACCTG 6281
Db 5599 TGCTTGGCAGCCCGCAACATCTCAAGGCGAGATTCGCGGGGCTCACAACACTGACCTG 5658
QY 6282 CGGGTCCCGCGCAAGGTTCCAGCGGCGGCGGCTTCTGTTGGCGGCGGAGTTCCCTGGG 6341
Db 5659 CGGGTCCCGCGCAAGGTTCCAGCGGCGGCGGCTTCTGTTGGCGGCGGAGTTCCCTGGG 5718

Qy	1590	CAAGTGGCATCAAAATCCGCATGGACATTGACGTGGTTCACAGGACCAATAAGATCAGG	1649
Db	1381	CAAGTGGCATCAAAATCCGCATGGACATTGACGTGGTTCACAGGACCAATAAGATCAGG	1440
Qy	1650	GACAGGTTTTGGGACCCCTGGCCACAGCCGCGAACCCCTGACCGACCTTGCGCTACGTGTGG	1709
Db	1441	GACAGGTTTTGGGACCCCTGGCCACAGCCGCGAACCCCTGACCGACCTTGCGCTACGTGTGG	1500
Qy	1710	GGCGGCTTCGTGTACCTGCAAGACCTGGTGAGCGGTGCAGCGGTTCGCGTGTCTCAGCGGC	1769
Db	1501	GGCGGCTTCGTGTACCTGCAAGACCTGGTGAGCGGTGCAGCGGTTCGCGTGTCTCAGCGGC	1560
Qy	1770	GCBAACCCCGGGCCGCTCTACCTGACAGATGCCCTTACCGTGTCTATGTGGACGAC	1829
Db	1561	GCBAACCCCGGGCCGCTCTACCTGACAGATGCCCTTACCGTGTCTATGTGGACGAC	1620
Qy	1830	GTGTTCTCTGGGTGTGCTGAGCCGCTGCTGCTGCTCTCTGCTGAGCTGGCTCTGATCTAC	1889
Db	1621	GTGTTCTCTGGGTGTGCTGAGCCGCTGCTGCTGCTCTCTGCTGAGCTGGCTCTGATCTAC	1580
Qy	1890	TCCGTGACATGACAGTGAAAGCCGTGGTGGGAGAAAGAGACGCGGCTGCGGGACACC	1949
Db	1681	TCCGTGACATGACAGTGAAAGCCGTGGTGGGAGAAAGAGACGCGGCTGCGGGACACC	1740
Qy	1950	ATGCGCGCCATGGGGCTCAGCGCGCGGTGCTCTGGCTAGGCTGGTTCTCAGCTGCCTC	2009
Db	1741	ATGCGCGCGGTGGGGCTCAGCGCGCGGTGCTCTGGCTAGGCTGGTTCTCAGCTGCCTC	1800
Qy	2010	GGGCGCTTCCTGCTCAGCGCGCGCTGCTGCTGCTCAAGCTGGGGGACATCCTC	2069
Db	1801	GGGCGCTTCCTGCTCAGCGCGCGCTGCTGCTGCTCAAGCTGGGGGACATCCTC	1860
Qy	2070	CCCTAGACCAACCGGCGGTGCTTCTCTGTTCTGGACGCTTCGCGGTGGCCACCGTG	2129
Db	1861	CCCTAGACCAACCGGCGGTGCTTCTCTGTTCTGGACGCTTCGCGGTGGCCACCGTG	1920
Qy	2130	ACCCAGAGCTTCCTGCTCAGCGCTCTCTCCGCGCCAACTGGCTGGCGCTGGCGC	2189
Db	1921	ACCCAGAGCTTCCTGCTCAGCGCTCTCTCCGCGCCAACTGGCTGGCGCTGGCGC	1980
Qy	2190	GGCTGGCCCTACTTCTCCCTCTACTGCTCCCTACTGCTGTGTGTGGCTGGCGGGACCGG	2249
Db	1981	GGCTGGCCCTACTTCTCCCTCTACTGCTCCCTACTGCTGTGTGTGGCTGGCGGGACCGG	2040
Qy	2250	CTGCCCCGGGTGGCCCGGTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2309
Db	2041	CTGCCCCGGGTGGCCCGGTGGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2100
Qy	2310	TGCGAGAGCTGCTGCTGAGGAGCAGGGCGAGGGCGGCGAGTGGGCAACGTGGGC	2369
Db	2101	TGCGAGAGCTGCTGCTGAGGAGCAGGGCGAGGGCGGCGAGTGGGCAACGTGGGC	2160
Qy	2370	ACCGGCTTACCGGACAGCTTCTCAGCTGGCCCAAGCTCTGCGCTTCTGCTGCTGAC	2429
Db	2161	ACCGGCTTACCGGACAGCTTCTCAGCTGGCCCAAGCTCTGCGCTTCTGCTGCTGAC	2220
Qy	2430	CGCGCGCTTACCGGCTTGGCACTGCTGTA CTTGGAAGCTGTGTGCTGCTGCTGCTGCTG	2489
Db	2221	CGCGCGCTTACCGGCTTGGCACTGCTGTA CTTGGAAGCTGTGTGCTGCTGCTGCTGCTG	2280
Qy	2490	ATCCCTGAA CCAATGGAAATTTCTTTTTCGAGGAGCTACTGCTGGGACCTTCGCGCCCC	2549
Db	2281	ATCCCTGAA CCAATGGAAATTTCTTTTTCGAGGAGCTACTGCTGGGACCTTCGCGCCCC	2340
Qy	2550	AAGAGTCCAGCCCTTGGCCCAACCCGCTGGA CCGAAGGTGTGCTGCTGCTGCTGCTG	2609
Db	2341	AAGAGTCCAGCCCTTGGCCCAACCCGCTGGA CCGAAGGTGTGCTGCTGCTGCTGCTG	2400
Qy	2610	CCCGGCTGAGTCTGCTGCTATCCGTTTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2669
Db	2401	CCCGGCTGAGTCTGCTGCTATCCGTTTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2460

Qy	2670	CAGCCAGCCCTGCGGGGCTCAGCTGGAATTCTACAGGGGCCACATCA CCGCTTCCTG	2729
Db	2461	CAGCCAGCCCTGCGGGGCTCAGCTGGAATTCTACAGGGGCCACATCA CCGCTTCCTG	2520
Qy	2730	GGCCA CAACGGGGCCGGCAAGACCA CACTGTTCATCTTGTAGTGGCTCTTCCACCC	2789
Db	2521	GGCCA CAACGGGGCCGGCAAGACCA CACTGTTCATCTTGTAGTGGCTCTTCCACCC	2580
Qy	2790	AGTGGTGGCTCTGCTTCTTCTTGGGGCAGAGCTCGCTCCAGCATGGCGGCCATCCGG	2849
Db	2581	AGTGGTGGCTCTGCTTCTTCTTGGGGCAGAGCTCGCTCCAGCATGGCGGCCATCCGG	2640
Qy	2850	CCCCACTCGGGCGTCTGTCTCAGTACAACTGCTGTTTGACATGTGTGACCCGTGGACGAG	2909
Db	2641	CCCCACTCGGGCGTCTGTCTCAGTACAACTGCTGTTTGACATGTGTGACCCGTGGACGAG	2700
Qy	2910	CAGCTCTGGTTCTATCGGGCGGCTGAAGGCTCTGAGTCCCGCTGTGTAGTGGCCCCGAGCAG	2969
Db	2701	CAGCTCTGGTTCTATCGGGCGGCTGAAGGCTCTGAGTCCCGCTGTGTAGTGGCCCCGAGCAG	2760
Qy	2970	GACCGTCTGCTGACAGATGTGGGGCTGGTCTCCAAAGAGAGTGTGACAGCTCGCCACCTC	3029
Db	2761	GACCGTCTGCTGACAGATGTGGGGCTGGTCTCCAAAGAGAGTGTGACAGCTCGCCACCTC	2820
Qy	3030	TCGTGTGGATGCAACGGAAGCTGTCCGTGGCCATTGCTTGTGGCGGCTCCCAAGTT	3089
Db	2821	TCGTGTGGATGCAACGGAAGCTGTCCGTGGCCATTGCTTGTGGCGGCTCCCAAGTT	2880
Qy	3090	GTTTATCCTGGACGAGCTACGGCTGCTGGTGGATTCCTGCTTCCCGCGCGGTATTTGGGAG	3149
Db	2881	GTTTATCCTGGACGAGCTACGGCTGCTGGTGGATTCCTGCTTCCCGCGCGGTATTTGGGAG	2940
Qy	3150	CTGCTGCTCAATACCGAGAGGTCCGACGCTGATCTCTCCACCCACCACTGGATGAG	3209
Db	2941	CTGCTGCTCAATACCGAGAGGTCCGACGCTGATCTCTCCACCCACCACTGGATGAG	3000
Qy	3210	GCAGAGCTGTGGGAGACCGTGTGGTGGTGGAGGTGGCGCTTGTGCTGCTGTGGC	3269
Db	3001	GCAGAGCTGTGGGAGACCGTGTGGTGGTGGAGGTGGCGCTTGTGCTGCTGTGGC	3060
Qy	3270	TCCCCACTCTTCTGCTGCGCTCACTGGGCTCCGGCTTACTGACTGCTGTGTGAAGGCC	3329
Db	3061	TCCCCACTCTTCTGCTGCGCTCACTGGGCTCCGGCTTACTGACTGCTGTGTGAAGGCC	3120
Qy	3330	CGCTGCTGCTGACCAACCAATGAGAGGCTGACCTGACATGAGAGGGAGTGTGGACAC	3389
Db	3121	CGCTGCTGCTGACCAACCAATGAGAGGCTGACCTGACATGAGAGGGAGTGTGGACAC	3180
Qy	3390	AGGCAAGAAAGAAAGATGGCAGCCAGGAGCAGAGTCCGCACTCTCTCAGCTGTGGCC	3449
Db	3181	AGGCAAGAAAGAAAGATGGCAGCCAGGAGCAGAGTCCGCACTCTCTCAGCTGTGGCC	3240
Qy	3450	CTGGTACAGCACTGGGTGCTCGGGGCA CCGCTGGTGGAGAGCTGCCACACGAGCTGGTG	3509
Db	3241	CTGGTACAGCACTGGGTGCTCGGGGCA CCGCTGGTGGAGAGCTGCCACACGAGCTGGTG	3300
Qy	3510	CTGGTCTGCTTACACAGGGTGCCTATGACGGGCTTCCGCACTCTTCCAGAGCTA	3569
Db	3301	CTGGTCTGCTTACACAGGGTGCCTATGACGGGCTTCCGCACTCTTCCAGAGCTA	3360
Qy	3570	GACACGCGGTGGCGAGCTGAGGCTCAGTGGCTACGGGATCTCCGACACAGCTCGAG	3629
Db	3361	GACACGCGGTGGCGAGCTGAGGCTCAGTGGCTACGGGATCTCCGACACAGCTCGAG	3420
Qy	3630	GAGATCTTCTGAAAGTGGTGGAGGAGTGTGCTGGGACACAGATATGGAGGATGGCAGC	3689
Db	3421	GAGATCTTCTGAAAGTGGTGGAGGAGTGTGCTGGGACACAGATATGGAGGATGGCAGC	3480
Qy	3690	TGCGGGGAGGACCTATGACAGGCATTTGCTGGCTTAGA CATAACCTCGGCTCAAGATG	3749
Db	3481	TGCGGGGAGGACCTATGACAGGCATTTGCTGGCTTAGA CATAACCTCGGCTCAAGATG	3540
Qy	3750	CCGCCACAGGAGCAGCGCTGGAGAAACGGGAAACAGCTGGGTGAGTGGTCCGCCACAGATG	3809

Db 3541 CCGCACAGGAGACAGCTGGAGAACGGGAAACAGCTGGGTGACCCCAAGACTGAC 3600
Qy 3810 CAGGGCTCTGGGCGAGAGCCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGAGAG 3869
Db 3601 CAGGGCTCTGGGCGAGAGCCGCTGGGCGGGGTACAGGGCTGGGCACTGACCCGCGAGAG 3660
Qy 3870 CTCGAGGCGCTGCTCTCAAGCGCTTTCTGCTGTCGCGCGAGAGCCGCGGGCGCTGTTTC 3929
Db 3661 CTCGAGGCGCTGCTCTCAAGCGCTTTCTGCTGTCGCGCGAGAGCCGCGGGCGCTGTTTC 3720
Qy 3930 GCCCAGATGCTGCTGCTGCTGCTCTTTGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3989
Db 3721 GCCCAGATGCTGCTGCTGCTGCTCTTTGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3780
Qy 3990 CCTCTCTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGTCAGGTGCTCAGGTG 4049
Db 3781 CCTCTCTTTCGGGCACTACCGGCTCTGCGGCTCAGTCCACCATGTCAGGTGCTCAGGTG 3840
Qy 4050 TCCTTCTTTCAGTAGGACGCGCCAGGGACCTTGACGCTGCGCGCTGCTCGAGGCGCTG 4109
Db 3841 TCCTTCTTTCAGTAGGACGCGCCAGGGACCTTGACGCTGCGCGCTGCTCGAGGCGCTG 3900
Qy 4110 CTGCAGGAGGAGGACTGGAGAGCCCGCAGTGACGATAGTCTCCACAGGTTCTCGGCA 4169
Db 3901 CTGCAGGAGGAGGACTGGAGAGCCCGCAGTGACGATAGTCTCCACAGGTTCTCGGCA 3960
Qy 4170 CCAGAAAGTTCTGCTGAAGTGGCAAGGTCTTGGCCAGTGGCACTGGACCCCAAGAGTCT 4229
Db 3961 CCAGAAAGTTCTGCTGAAGTGGCCAGGTTCTTGGCCAGTGGCACTGGACCCCAAGAGTCT 4020
Qy 4230 CCATCCCGAGCTGCGAGTGTAGCCAGCCCGGTGCGCGGCTGCTGCGCGAGTGCCTCCG 4289
Db 4021 CCATCCCGAGCTGCGAGTGTAGCCAGCCCGGTGCGCGGCTGCTGCGCGAGTGCCTCCG 4080
Qy 4290 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4349
Db 4081 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4140
Qy 4350 AACCTGACAGCCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4409
Db 4141 AACCTGACAGCCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4200
Qy 4410 CAGGCTGGAAGACTGAAGAGTGGTGAATGAGGTGAGTACGAGGCTTCTCGCTGGGG 4469
Db 4201 CAGGCTGGAAGACTGAAGAGTGGTGAATGAGGTGAGTACGAGGCTTCTCGCTGGGG 4260
Qy 4470 GCGCGAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4529
Db 4261 GCGCGAGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4320
Qy 4530 GCGCTGCTGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4589
Db 4321 GCGCTGCTGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4380
Qy 4590 TGGGCTCACAGCTGAGTCTCAGGACAGTCTCAGAGTCTGCTGCTGCTGCTGCTGCTGCTG 4649
Db 4381 TGGGCTCACAGCTGAGTCTCAGGACAGTCTCAGAGTCTGCTGCTGCTGCTGCTGCTGCTG 4440
Qy 4650 CACTCCATGCTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCTCGGTGCTCACTGCTGCC 4709
Db 4441 CACTCCATGCTGGCTTTGTCAACCGAGCCAGCAACGCAATCTCTCGGTGCTCACTGCTGCC 4500
Qy 4710 CCAGGCGCGGCGCGCCAGCCACAGCATCACCACCTCACCACCTCACCACCTCACCACCTCACC 4769
Db 4501 CCAAGCGCGGCGCGCCAGCCACAGCATCACCACCTCACCACCTCACCACCTCACCACCTCACC 4560
Qy 4770 AAGGAGCAGCTGCTGAGGCTGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4829
Db 4561 AAGGAGCAGCTGCTGAGGCTGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
Qy 4830 TGTGTGCTTTTGCCATGCTCTTTGTCGCGGCGCAGCTTCACTCTTGTGCTCTCACTGAGGAG 4889

Db 4621 TGTGTGCTTTTGCCATGCTCTTTGTCGCGGCGCAGCTTCACTCTTGTCTCTCATTTGAGGAG 4680
Qy 4890 CGAGTCAACCGAGCCAAAGCACTGAGCTCATGCGGGGCTGTCGCCCACTCTACTGAG 4949
Db 4681 CGAGTCAACCGAGCCAAAGCACTGAGCTCATGCGGGGCTGTCGCCCACTCTACTGAG 4740
Qy 4950 CTTGGCAACTTTCTCTGGGACATGTAACCTTCTGTTGGTCCAGCATGATGCTGCTGCTGCTG 5009
Db 4741 CTTGGCAACTTTCTCTGGGACATGTAACCTTCTGTTGGTCCAGCATGATGCTGCTGCTGCTG 4800
Qy 5010 ATCTTTCTGGGCTTCCAGCAGAGGCAATGTCGCGCTGCAACCTGCTGCTGCTGCTGCTG 5069
Db 4801 ATCTTTCTGGGCTTCCAGCAGAGGCAATGTCGCGCTGCAACCTGCTGCTGCTGCTGCTG 4860
Qy 5070 CTGTGTCTACTGTAATGCTGCTGATCACCGCTCATGTACCCAGCTCTCTCTCTCTCT 5129
Db 4861 CTGTGTCTACTGTAATGCTGCTGATCACCGCTCATGTACCCAGCTCTCTCTCTCTCTCT 4920
Qy 5130 TTCTCCGTGCGCAGCAGACGCTATGTTGCTGCTACCTGTCATATAAACCCTCTTTATTGGCATC 5189
Db 4921 TTCTCCGTGCGCAGCAGACGCTATGTTGCTGCTACCTGTCATATAAACCCTCTTTATTGGCATC 4980
Qy 5190 AATGGAAGCATGGGCACTTTGCTGCTGAGCTCTTCTGATCAGAACTGAGAGAGTG 5249
Db 4981 AATGGAAGCATGGGCACTTTGCTGCTGAGCTCTTCTGATCAGAACTGAGAGAGTG 5040
Qy 5250 AGCCGGATCTTTGAACAGGCTCTCTTATCTTCCCACTTCTGCTTGGGCGGGGGCTC 5309
Db 5041 AGCCGGATCTTTGAACAGGCTCTTCTTATCTTCCCACTTCTGCTTGGGCGGGGGCTC 5100
Qy 5310 ATTGACATGCTGCGGAACAGGCTGCTGATGCTTGTGAGCGCTTGGGAGAGAGGCGAG 5369
Db 5101 ATTGACATGCTGCGGAACAGGCTGCTGATGCTTGTGAGCGCTTGGGAGAGAGGCGAG 5160
Qy 5370 TTCCAGTCAACCCCTGCGCTGGGAGGCTGTCGCAAGAACTTCTTGGCCATGCTGATACAG 5429
Db 5161 TTCCAGTCAACCCCTGCGCTGGGAGGCTGTCGCAAGAACTTCTTGGCCATGCTGATACAG 5220
Qy 5430 GGGCCCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5489
Db 5221 GGGCCCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5280
Qy 5490 CCCAGGTGAGGTCTCTGCCACTCTCTGGGAGGAGGAGGAGGATGTAGCCCGTGAACGG 5549
Db 5281 CCCAGGTGAGGTCTCTGCCACTCTCTGGGAGGAGGAGGAGGATGTAGCCCGTGAACGG 5340
Qy 5550 GAGCGGTGCTTCAAGGAGCCACCCAGGGGATGTTGTTGGTCTGAGGAACTTGAACCAAG 5609
Db 5341 GAGCGGTGCTTCAAGGAGCCACCCAGGGGATGTTGTTGGTCTGAGGAACTTGAACCAAG 5400
Qy 5610 GTATACCGTGGGAGGAGGATGCCAGCTGTTGACCGCTTGTGCTTGGGATTTCCCGCTGGT 5669
Db 5401 GTATACCGTGGGAGGAGGATGCCAGCTGTTGACCGCTTGTGCTTGGGATTTCCCGCTGGT 5460
Qy 5670 GAG 5672
Db 5461 GAG 5463

Search completed: December 30, 2004, 13:41:53
Job time : 3248 secs

This Page Blank (uspto)

Result No.	Query			Description		
	Score	Match	Length	ID	DB	
C	1	1628.8	24.2	6606	9	AY405471 Homo sapi
	2	1628.8	23.9	6600	9	AY405473 Mus muscu
	3	1559.4	22.9	6606	9	AY405472 Pan trogl
	4	1150.8	16.9	2110	6	CB606356 AGMNNUC:S
	5	904.4	13.3	4783	3	BK051920 Mus muscu
	6	828.8	12.2	933	5	BQ959013 AGENCOURT
	7	820.2	12.1	3759	3	CR627391 Homo sapi
	8	805.8	11.8	3605	3	AK052916 Mus muscu
	9	799	11.7	958	5	BQ955522 AGENCOURT
	10	760.6	11.2	1934	3	BC036366 Mus muscu
C	11	751.8	11.0	914	5	BUS28031 AGENCOURT
	12	748.2	11.0	969	5	BUB46021 AGENCOURT
	13	733.2	10.8	884	4	BG775870 602650251
	14	717.4	10.5	749	6	CD629699 55052443J
	15	704.8	10.4	730	6	CD629709 55058387J
	16	700.6	10.3	941	2	BE793677 601590063
	17	684.4	10.1	880	4	BG674537 602620581
	18	680.2	10.0	1415	3	BC038521 Mus muscu
	19	678.2	10.0	705	6	CD629715 56032448J
	20	675.6	9.9	714	6	CD629703 55058323J
C	21	673	9.9	681	6	CAS14076 UI-CF-FNO
	22	670.4	9.9	782	4	BG775975 602650251
	23	655.4	9.6	707	6	CD629704 55058371H
	24	653.8	9.6	668	6	CD629707 55058379J

Db 862 GACAACTCTACAACTCTTACTGCAATGATTTGATGAAGAAATTTGGAGTCTAGTCTCTCTT 921
Qy 1047 TCCGCCCTGTCTGGAGACGCTGAAGCCTCTGATCTCGGGAAGTACTTCTTGGACCA 1106
Db 922 TCCGCCATTATCTGAAAGCTCTGAAGCGCTGCTCTGTTGGAAGATCTCTGTATACACT 981
Qy 1107 GATACACTTTTACCCGGAAGCTATGCCCCAGGTGAACGGACCTTCGAGGAGCTCAC 1166
Db 982 GACACTCCAGCAACAAGCAAGGTCTATGGCTAGGTGAACAGACTTCCAGGAACCTGGCT 1041
Qy 1167 CTGCTGAGGATGTCGGGAGGTCTGGAGATGCTGGACCCCGGATCTTCACTTCTATG 1226
Db 1042 GTGTTCCATGATCTGGAAGCAATGTGGAGAACTCAGCCCCCAAGATCTGACCTTCATG 1101
Qy 1227 AACGACAGTTTCAATGTGGCAATGCTCAGCGGCTCTCGAGATGACAGGA----- 1276
Db 1102 GAGAAACAGCAAGAAATGGAACCTTGTCCGATGCTGTTGGACAGCAGGCAATGACCAC 1161
Qy 1277 -----TGAAGGAAGAGGAGCCAGCCAGACTGGAGCCGGACCAATGG 1320
Db 1162 TTTTGGAAACAGCAGTTGGATGGCTTAGATTGGACAGCCCAAGACATCGTGGCGTTTTG 1221
Qy 1321 ----AGGCCCTGCGATCTTCTGACCTCGGAGCGGTGCTACAGCTGGCAGGACGCA 1376
Db 1222 GCCAAGCACCCAGAGGATGCCAGTCCAGTAATGGTTCTGTGTACACTGGAGAGACT 1281
Qy 1377 CACGCTGATGTGGGCACTGTGGGCACGCTGGGCGAGTGACGGAGTGCCTGTCTTG 1436
Db 1282 TTCAACGAGACTAACACAGGCAATCCGAGCAATATCTCGTTCTATGGAGTGTCAACCTG 1341
Qy 1437 GACAACTGGAGGCGGACCTCAGAGGACGCCGTGTGTCGGGCGCTGCACTGCTC 1496
Db 1342 AACAACTAGAACCCATAGCAACAGAACTGTGGCTCATCAACAAAGTCCATGGAGCTGTG 1401
Qy 1497 GCGGAACATCGATTCTGGGCGCGCTGCTCTTCTTGGGACCTGAGGACTCTTCAGACCC 1556
Db 1402 GATGAGAGGAATTTCTGGGCTGGTATTTGTGTTCACTGGAAT-----ACTCCAGGAGC 1455
Qy 1557 ACAGACACCCAAACCCAGACTGGGCGCGGCGCACGTCGGGCAATCAAAATCCGATGGAC 1616
Db 1456 ATTGAGC-----TGCCCCCATCATGTCAAGTACAAGATCCGAATGGAC 1497
Qy 1617 ATTGACGTGTCAGAGCAACATAAGATCAGGACAGGTTTGGGACCTTGGGCCAGCC 1676
Db 1498 ATTGACAAATGGAGAGGACAAATAAAATCAAGATGNNNNNNNNNNNNNNNNNNNN 1557
Qy 1677 GCGGACCCCTGACCGGACTCGCTAGCTGTGGGCGGCTTCTGTGTACTCTGCAAGACTG 1736
Db 1558 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1617
Qy 1737 GTGGAGCGTGAGCCGCTCCGCTGCTCAGCGGCGCAACCCCGGGCGGCTCTACTG 1796
Db 1618 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1677
Qy 1797 CAGCAGATGCCCTATCCGTGCTATGTGAGACGCTGTTCTCGTGGTGCTGAGCCGGTCG 1856
Db 1678 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1737
Qy 1857 CTGCGGCTTTCTGACGCTGGCTGATCTACTCCGTGACACTCAGCAGTGAAGCCGCTG 1916
Db 1738 ATGCCCTCTTCAAGACGCTGGCTGATTTACTCAGTGGCTGTGATCATCAAGGGCATC 1797
Qy 1917 GTGGGAGAGAGAGACCGGCTCGGGACACCATGTCGCGCATGTGGGCTCAGCCGCGG 1976
Db 1798 GTGTATGAGAGGAGGACCGCTGAAAGAGACCATGCGGATCATGGGCTTGAACAGC 1857
Qy 1977 GTGCTGTGGCTAGCTGTCTCAGCTGCTCGGCGCTTCTGCTCAGCGCGCGCTG 2036
Db 1858 ATCCTCTGTGTTAGCTGTGCTATTAGTACCTCATCTCTTCTGTGAGCGCTGGGCTG 1917
Qy 2037 CTGTTCTGTGTGCTCAAGCTGGGAGCATCTCCCTCCTACAGCCACCGCGGCTGCTTTC 2096
Db 1918 CTAGTGTGTCATCTGAAATAGGAAACCTGCTGCCCTACAGTATCCAGCGTGTGTTT 1977

Qy 2097 CTGTTCTTGGAGCCTTTCGGGTGGCCACGGTGACCCAGAGCTTCTGCTCAGGCGCTTC 2156
Db 1978 GTCTTCTGTCCGCTTGTGCTGTGGTGACAACTCTGCAGTGTCTTCTGATTAGCACACT 2037
Qy 2157 TTCTCCGGCGCAACCTGGCTCGGCTCGGGGCTGSCCTACTTCTCCCTTACTCTG 2216
Db 2038 TTCTCAGAGCCAACTTGGCAGAGCTGTGGGGCATCATCTACTTTCAGCTGTACTCTG 2097
Qy 2217 CCCTAGCTGTGTGTGGCTTGGCGGACCCGGCTGCCCCGGGTGGCCCGCTGGCGCG 2276
Db 2098 CCCTAGCTGTGTGTGGCATGSCAGGACTACGTGGGCTTACACTCAAGATCTTCTGCT 2157
Qy 2277 AGCTGTGTGCGCGCTGGCTTCCGCTTCCGCTGCGAGAGCTTGGCTGTCTGTGGAGAG 2336
Db 2158 AGCTGTGTGTCTCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTCCTCTTTTGGAG 2217
Qy 2337 CAGGCGAGGCGCGAGTGGCAACAGTGGGCACCCGGCT--ACGGCAGACGCTTTC 2393
Db 2218 CAGGCAATTGGAGTGCAGTGGGAAACCTGTTTGGAGTCTCTGTGGAGGAAGTGGCTTC 2277
Qy 2394 AGCTGGCCCAAGTCTCTGGCTTCTGTCTGTGAGCGCGCTCTACGGCCTTCGCCACC 2453
Db 2278 AATCTCACCACTTTCGGTCTCCATGATGCTGTTTGACACCTTCTCTATGGGTGATGACC 2337
Qy 2454 TGGTACTGGAGCTGTGTGCCAGGCGAGTACGGGATCCCTGAAACCATGGAATTTTCT 2513
Db 2338 TGGTACTTGAAGCTGTCTTTCAGGCCAGTACGGAATTTCCAGGCCCTGATTTTCT 2397
Qy 2514 TTTTGGAGGAGTACTGTGTGGGACCTCGGCCCCCAAGAGTCCAGCCCTTCGCCCAACC 2573
Db 2398 TGCACCAAGTCTTCTGTTTGGCGAGGAAAGTGTAGAGAGGACCACTTGTTCACAC 2457
Qy 2574 CCGCTG---GACCCAAAGTGTGTGTAGAGAGGACCCGCCGCTGAGTCTCTGGCGTA 2630
Db 2458 CAGAAGAGAATATCAGAAATCTGCATGGAGAGGAAACCCACCACTTGAAGCTGGGCGTG 2517
Qy 2631 TCCGTTGCGAGCTGGAGAGCCGCTTCTGTGGAAGCCCGAGCCAGCCCTTCGGGGGCTC 2690
Db 2518 TCCATTCAGAACTGTGTAAAGTCTACCGAGATGGGATGAAGTGGCTGTGATGGCGTG 2577
Qy 2691 AGCTGGACTTCTACACGGGCCACATCACCGCCTTCTGGGCCACAAACCGGGCGGCAAG 2750
Db 2578 GCACTGAATTTTATGAGGCGCAGATCACTCTTCTTGGGCCACAAATGAGCGGGGAG 2637
Qy 2751 ACCACCACTGTCTCAATCTTGTAGTGGCTTCTCCACCCAGTGGTGGCTCTGCTTCTCATC 2810
Db 2638 ACACCAACATGTCAATCTCTGACCGGCTTGTTCCTCCCGGACCTCGGGGACACCGCTACATC 2697
Qy 2811 CTGGGCCACGAGCTCCGCTCCAGCATGGCCGCTTCCGGCCCGCCCTTCGGGCGCTGCTCCT 2870
Db 2698 CTGGGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTTGGGGGCTGTGCC 2757
Qy 2871 CAGTACAACTGTCTGTTGACATGCTGACCGTGGAGAGACGCTGTGTTCTATGGGCGG 2930
Db 2758 CAGCATAACTGTCTGTTGACATGCTGACTGTGGAAGAACACATCTGTTCTATGCCCGC 2817
Qy 2931 CTGAAGGCTCTGAGTCCGCTGTAGTGGCCCCGAGCAGACCGCTGTGCTGAGGATGTG 2990
Db 2818 TTGAAGGGCTCTCTGAGAAGCAGTGAAGCGGAGATGAGAGAGATGGCCCTGGATGTT 2877
Qy 2991 GGCTGTCTC---CAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATGCAACGG 3047
Db 2878 GGTTCCTCATCAAGCAAGCTGAAAAAGCAAAACAGCCAGCTGTGAGTGAATGAGAGA 2937
Qy 3048 AAGCTGTCCGCTGGCCATTGCTTTTGGGCGGCTCCCAAGTTGTATCTCTGGAGAGCCT 3107
Db 2938 AAGCTATCTGTGGCTTGGCTTTTTCGGGGGATCTAAGGTTGTCTTCTGGATGAACCC 2997
Qy 3108 ACAGCTGTGGGATCTGCTTCCCGCCCGGATTTTGGAGAGCTGTGCTCAATACCGA 3167
Db 2998 ACAGCTGTGTGACCCCTTACTCCCGCAGGSGAATATGGAGCTGTGCTGNAATACCGA 3057


```
Db 5198 ACAATAAGCTCAATGACATCAATGACATCCTGAAAGTCTGTGTCTTATCTTTCCACATTT 5257
Qy 5290 TCTGTCTGGCGCGGGGCTCATTTGACATGTTGGTGGGAACAGAGCCATGCTGATGCTTTTG 5349
Db 5258 TTTGCTTGGGCGAGGGCTCATCGACATGTTGAAGAACAGAGCCATGGCCGATGCCCTGG 5317
Qy 5350 AGCGTCTGGGAGACAGGAGTTTCCAGTCAACCCCTCGCTGGGAGTGTTGGCAAGAAC 5409
Db 5318 AGAGGTTTGGGAGAACCGCTTCTCTCGCTCTCTGGGACTTGGTAGGACGGAACC 5377
Qy 5410 TCTTGGCCATGTTGATACAGGGGCCCCCTTCTTCTCTTACACTACTGCTGAGACAC 5469
Db 5378 TTTTGGCCATGTCGCGTGAAGGGGTGTCTTCTCTTACTACTGTTCTGATCCAGTACA 5437
Qy 5470 GAAGCCAACTCTCTGCCACAGCCAGGGTGAAGTCTCTGCACTCTCTGGAGAGGAGGACG 5529
Db 5438 GATTTTTTTCATCAGGCCGACACTGTAAGGCGAAGCTTCTCTTGAATGACGAGGATG 5497
Qy 5530 AGGATGTAGCCCGTGAACCGGGAGCGGGTGTCCAAAGAGACCCAGGGGGATGTGTTGG 5589
Db 5498 AGGATGTAGGGCGGAGAGACAGAGGATTTCTGACGCTGGCGGACAGAAATGACATCCTAG 5557
Qy 5590 TGCTCAGGAACTTGACCAAGGTATACCGTGGGACAGGATGCCACTGTTGACCGCTTGT 5649
Db 5558 AGATCAAGGAACTGACCAAGATCTATAGGAGGAAGCGAAGCCTGCAGTCCGACAGATCT 5617
Qy 5650 GCTCGGGATTCCTCCCTGCTGATGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5709
Db 5618 GCATCGGATTCCTCCGAGAGTCTTGGACTCTGGAGTTAACGGAGTGGGAAGT 5677
Qy 5710 COTCACAAGTTTGGCATGTTGACGGGGACACATTTGCCAGCAGGCGGCGAGGCTGTGCTGG 5769
Db 5678 CAACAACATTTCAAGATGCTGCTGGAGACACCCCTGTGACAGAGGGATGGTTCCT-- 5735
Qy 5770 CAGGCCACAGCTGGCGCGGAAACCCAGTGTGCTGCCACTCAGCATGAGGATGCTGCCCTC 5829
Db 5736 ----TAAACAATCTTATCAATATCCATGAATACACAGAAATGGGCTACTGCCCTC 5791
Qy 5830 AATCCGATGCCATCTTTGAGTGTCTGAGCGCGCGGACACTGAGCTGCTGGCGCC 5889
Db 5792 AGTTTGACGCATCAACAGCTGCTGAGCGGAAAGAGACATGTGAGTCTTTGGCCCTCC 5851
Qy 5890 TGCGGCTGTCCCGAGGCGCCAGGTTGCCAGACCGCTGGCTCAGGCGCTGGCGGTCTGG 5949
Db 5852 TGAGGGAGTCCAGAAAGGAGTTGGCAAGTTGGTGAATGGCAATTCGCAACTGG 5911
Qy 5950 GACTCTCATGTGACGACAGCGGCTGAGGACCTACAGCGGAGGGAACAAACGCAAGC 6009
Db 5912 GCCTGTTAAAGTATGGAGAAAAATATGCCAGTAACTACAGTGGCGGCAACAAACGAAAGC 5971
Qy 6010 TGGGACCGGCTGGCGCTGTTGGGACCCAGCGCTGTTCTTGTGACGAGCGGACCA 6069
Db 5972 TCTCAGACCATGCTTGTATGTCGGGACCTCTCTGCTGTTCTGATGAACCAACCA 6031
Qy 6070 CAGGATCGACCCAGCGCGGCGCTTCTTTTGGAAACAGCCTTTTGGCGGTGGCGGG 6129
Db 6032 CAGGATCGAACCTTAAAGCCCGGAGATCTTGTGGAATTTGCCCTAAGCATTTGCAAGG 6091
Qy 6130 AGGGCCGTTCAATGATGCTCACTCTCCATAGCATGAGGAGTGAAGCGCTCTGCTCGC 6189
Db 6092 AGGGGAGATCTGTAGTCTTACATCTCATATGATGTAAGAAATGTGAAGCTCTTTGTACA 6151
Qy 6190 GCCTAGCCATCATGTGTAATGGGCGGTTCCGCTGCTGGGAGCGCCGAAACATCTCAAGG 6249
Db 6152 GGATGGCCATTAATGTGTAATGGAAGGTTTCAAGTGTCTTGGAGTGTCCAAATCTGAAA 6211
Qy 6250 GCAGATTGCGCGGGTCAACACTGACCTTGGCGGTGCTCCCGCGCAAGTCCC----- 6303
Db 6212 ACAGTTTGGAGATGTTTATACATAGTTGTACGAATAGAGGCTCCAAACCTGACCTGA 6271
Qy 6304 AGCCGGCAGCGGCTTCTGTTGGCGGCGGAGTTCCCTGGGTGCGAGCTGCGCGAGGCACATG 6363
Db 6272 AGCCTGTCCAGGAGTTCTTTGGACTTGTTCGGGAAAGTGTCTTAAAGAGAAACATC 6331
```

```
Qy 6364 GAGGCGGCTGCGCTTCCAGCTGCCGCGGAGGCGCTGCCCTGGCGCGCTTTTG 6423
Db 6332 GAACATGCTTTCAGTACCAGCTTCCATCTCTTGTCTAT--CTTACGAGGATATTCA 6388
Qy 6424 GAGAGCTGGCGGTGACGCGCAGAGCAGCGGTGGAGGACTTTTCCGTGAGCCAGACA 6483
Db 6389 GCATCCTCTCCAGAGCAAAAAGGCACTCCACATAGAAGACTACTCTCTCTCAGACAA 6448
Qy 6484 TGCTGGAGGAGTATTCTTGTACTTCTCCAGGACCCAGGGAAGACGAGGACCCGAG 6543
Db 6449 CACTTGACCAAGTATTTGTGAACCTTGCCAAAGGACCAAAAGTATGATGACCACTTTAAAGG 6508
Qy 6544 AGCAG 6548
Db 6509 ACCTG 6513

RESULT 3
LOCUS AY405472
DEFINITION Pan troglodytes ABCA1 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY405472
VERSION AY405472.1 GI:39761446
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 6606)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6606)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..6606
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>6606
/gene="ABCA1"
/locus_tag="HCM2235"

gene
Query Match 22.9%; Score 1559.4; DB 9; Length 6606;
Best Local Similarity 52.7%; Pred. No. 5.2e-310;
Matches 3002; Conservative 0; Mismatches 2539; Indels 150; Gaps 14;

Qy 979 CCCTGCCAGACAGAGCCTGAGCCCGCGCTGCTCGGAGCTGATGGAGCCCTGGACAGCC 1038
Db 854 CTTNNNNNNNNNNNNNNNNNNCTCTTACTGTAATGATTTTGTGAAGAAATTTGGAGTCTA 913
Qy 1039 ACCCGCTGTCCTCGCTCTGGAGAGCCCTGAGCCTCTGATCTCTGGGAGCTACTCT 1098
Db 914 GTCTCTTTCCCGCATTTATCTGGAAAGCTCTGAAGCCGCTCTGTTGGGAGATCCTGT 973
Qy 1099 TTGACACAGATACACCTTTTATCCCGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGG 1158
```


Qy	3280	TCCTGGCCGCTCACTCGGCTCCGGCTCTACCTGACGCTGGTGAAGCCCGCGCTGCCCC	3333
Db	3170	TTCTGAAGAAACAGCTGGGAAACAGGCTACTACCTGACCTCGTCAAGAAAGATGTGGAAT	3229
Qy	3340	TGACCAACAAT-----GAGAAGGCTGACA	3363
Db	3230	CCTCCCTCAGTTCTCTGCAGAAAACGTAGTAGCACTGTGTCAATCCTGAAANAAGAGACA	3289
Qy	3364	CTGACATGGAGGCGAGTGTGGACACCAAGGAGGAAAGAAAGATGGCAGCCAGGGCAGCA	3423
Db	3290	GTGTTTCTCAGAGCAGTTCTGATGCTGCGCTGGCGACGACCATGATGAGTGAACACGCTGA	3349
Qy	3424	GAGTGGCAGCTCTCAGCTGCTGGCCCTGTGTACAGCACTGGGTGCCCGGGGCAACGCTGG	3483
Db	3350	CCATCGATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTCTCTGAAGCCCGCTGG	3409
Qy	3484	TGAGGAGCTGCCACACAGCTGGTGTCTGCCCTACACGGGTGCCCATGACGGCA	3543
Db	3410	TGGAAAGACATAGGGCATGAGCTGACTATGTGTCTGCCATATGAAGCTCTTAAGAGAGGGG	3469
Qy	3544	GCTTTCGCCACACTCTCCGAGAGCTAGACACGGGGCTGGCGAGCTGAGGCTCACTGGCT	3603
Db	3470	CTTTTGTGAACCTCTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTT	3529
Qy	3604	ACGGGATCTCCGACACAGAGCTCGAGAGATCTCTGTAAGGTGGTGAAGGATGTGCTG	3663
Db	3530	ATGGCATCTCAGAGACGACCTCGGAAGAAATATTTCTCAAGGTGGCCGAAGAGAGTGGG	3589
Qy	3664	CGGACACAGATATGGAGGATGGAGCTGCGGGCAGCACCTATGCANAGGCATTTGCTGGCC	3723
Db	3590	TGGATGCTGAGACCTCAGATGGTACCTTTCGACGAAAGACGACGCGGCGCTTCGGGG	3649
Qy	3724	TAGACGTAAACCTTCGGGCTCAAGATGCGGCCACAGGAGACAGCGCTGGAGAACGGGAAC	3783
Db	3650	ACAAGCAGAGCTGTCTTCGCCCGTTCTACTGAAGATGATGCTGTGTATCCAAATGATTCG	3709
Qy	3784	CAGCTGGGTGAG--CCCCAGAGACTGACAGGGCTCTGGGCCAGACGCGCTGG-----	3835
Db	3710	ACGTAGACCAGAAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAGGGCTCT	3769
Qy	3836	CCGGGTACAGGGCTGGGCACTGACCGCAGCAGCTCAGGGCCCTGCTTCTCAAGCGCT	3894
Db	3770	ACCAGGTGAAGGCTGGAATCTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGAC	3829
Qy	3895	TTCTGCTTGCCCGCGCAGCGCGGCTGTTGCGCCAGATCGTGTCTGCCCTGCCCTCT	3954
Db	3830	TGCTAATTCGACAGCGAGTTCGAAAGGATTTTTTGTCTCAGATGTCTTGCCAGCTGTGT	3889
Qy	3955	TTGTGGGCTGGGCCCTGTGTTTCAAGCTCATGTGCTCTCTTCGGGCACTACCCGGCTC	4014
Db	3890	TTGTCTGCAATGGCCCTGTGTTTCAAGCTGATCGTGCACCCCTTTGGCAAGTACCCCGAC	3949
Qy	4015	TGGGGCTCAGTCCCAACCATGTAAGGTGCTCAGTGTCTTCTTCAGTAGGAGCGCCCCAG	4074
Db	3950	TGGAACCTTCAGCCCTGGAATGTACAACAGCAAGTACACATTTGTTCAGNNNNNNNNNN	4009
Qy	4075	GGGACCCCTGGAAGCTGCCCGCTGCTGAGCGCTGTGTCAGGAGGCAAGACTGGAGGAGC	4134
Db	4010	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	4069
Qy	4135	CCCCAGTGCAATAGCTCCACAGGTTCTCGGCACCAAGAGTCTCTGTGTAAGTGGCCA	4194
Db	4070	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	4129
Qy	4195	AGTCTTGGCCAGTGGAACCTGACCCAGAGTCTCAATCCCAAGCTGCGAGTGTAGCC	4254
Db	4130	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	4189
Qy	4255	AGCCGTGCGCGCGCTGCTCCCGACTGCTCCCGCTGCGAGCTGTGTGTGCTCCCTCCGC	4314
Db	4190	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	4249
Qy	4315	CCGAGGAGTGAACCG-----CTCTGGGGAAGTGG	4344

Db 565 GCGAGCCTGGCCACAGAGAGTAACCAACCGCTTAAGCAA-----GGGTCAAGTA 615
Qy 679 CGGAGCTGCTGACGTCACTGCTGCCGACGGAATCCCTGGGGTTGGCACTGGGCCAAGCCC 738
Db 616 CTGAGCTCTTGGAGAAGATCTCTGCAACCGGGCATCTCTGGAGACCCTGTGGGTCAAGCCC 675
Qy 739 AGGAGCCCTTGCACAGCTTGTGGAGGCGCTGAGGACCTGGCCACAGAGCTCTTGGGCG 798
Db 676 AGGATTCATGAAAGTTCTCAAGATGCTACAGAACTGTTGCCACAGAGCTCTTGAAGC 735
Qy 799 TGGCAGGCTGGTGGAGCTTGGGCACTGCTGCAGAGACCCCGAGGGACCAAGCGCCCCC 858
Db 736 TACCAGGCTTGTGAACTCGGGCTTGTGTCGGAGGCGCCGAGGGTCAGCTGGCTCTC 795
Qy 859 TGGAGTGTCTGACAGGCCCCCTCTGAGTGTGAGGGACCTAGCAGCAGAGTGGGCCCT 918
Db 796 TGGAGCTGATTTGGAGGCCCCCTCTGCACTACCAAGGGGCCAGCAGTCCAGGGGGTCTGT 855
Qy 919 CCTCAACTGGTACGAGGCTAGTGACCTGATGGAGCTGTGGGGCAGAGCCAGAAATCCG 978
Db 856 CCTCAACTGGTACGAGGACCAACAGATCAAGCAATTCATGGGGCTGAGCTGGGCCCA 915
Qy 979 CCTGCCAGACAGCAGCTGAGCCCGCTGCTCGGAGCTGATTTGGAGCCCTTGGACAGCC 1038
Db 916 CCTGCTGACAGCAGTCTCAGCCCGCTGCTGAGTTCGTGGGGGCACTGGATGACC 975
Qy 1039 ACCGCTGTCCGCTGCTTGGAGACGCTGAGCCCTCTGATCTCGGGAAGCTACTCT 1098
Db 976 ACCCTGTGTCTGGCTGCTTGGAGACGCTGAGCCCTGATCTCTCGGGAATTTCTCT 1035
Qy 1099 TTGACACAGATACACTTTTACCCGGAAGCTCATGGCCAGGTGAACCGGACCTTCGAGG 1158
Db 1036 TTGACCTTGACACAACTTCACTAGAAACTCATGGCTCAGGTGAACAGACCTTTGAGG 1095
Qy 1159 AGCTCACCTGCTGAGGATGTCGGGAGGTGTGGGAGATCTGGGACCCCGGATCTTCA 1218
Db 1096 AGCTGGCTCTGTGAGGAGCTTACACGAACTCTGGGGAGTACTGGGACCCAGATCTTCA 1155
Qy 1219 CTTTCATGACGACAGTTCCATGTGGCCATGCTCAGCGGCTCTGAGATGAGGATG 1278
Db 1156 ACTTCATGAAAGACAGTAAACGTTGGCCATGCTTCAGAACTGCTGGATGAGGAGCA 1215
Qy 1279 AAGGAAGAAGCAGCCAGACCTGGAGCGGAGCACATGAGGCGCTCGGATCTTTC 1338
Db 1216 CAGGTTGACAGCAGCAGACACCAAGGCCAGACAGTGTGAGGCTATCAGAGCTTTC 1275
Qy 1339 TGGACCTTGGAGCGGTGGCTACAGCTGGCAGGACGACACGCTGATGTGGGCGACCTGG 1398
Db 1276 TGGATCTTAGTAGGGTTCGATACAACTGGCAGGAGGCTCATGCAGACATGGGCGCTTG 1335
Qy 1399 TGGGACGCTGGGCGGAGTGAGGAGTGGCTGTCTTGGACAGCTGAGGCGGACCT 1458
Db 1336 CAGAAATCTGGGCGCAATCTTGGAGTGTGTCTCGGACAGCTGGAGGCTGTGCCCT 1395
Qy 1459 CAGAGCAGCTGTGTGCGGGCCCTGCAACTGCTCGGAAACATGATTTGGGCG 1518
Db 1396 CAGAGAAAGCTTGTGTCCGTGGCCCTGGAGCTGTGGGTGAGCGCCCTCTTGGGAG 1455
Qy 1519 CGGTGCTTCTTGGAGCTGAGGACTCTTTCAGACCCACAGACCAACCCCA---G 1575
Db 1456 GCATCGCTTCTTGGAGCCAGAGCACCCTTGGACTCATCTGAAACCGCACTCCAAACA 1515
Qy 1576 ACCTGGGCGCCGCGCACTGGCGCATCAAAATCCGATGGCATGAGCTGTGTCAGGAGA 1635
Db 1516 CCACGGGTCTTGGCCACCTACAGAGTCAAGATTCGAATTTGATGATGTCAAGGA 1575
Qy 1636 CCAATAAGATCAGGACAGGTTTTGGGACCTTGGCCAGCGCCGACCCCTGACCCGACC 1695
Db 1576 CCAATAAGATCAGGACAGGTTTTGGGACCCAGAGTCCGTGAGCAGATCCCTCATGGACC 1635
Qy 1696 TGGCTAGCTGTGGGCGGCTTCTGTGTACCTGCAAGACCTGTGTGAGGCGTGCAGCGCTCC 1755

Db 1636 TTGCTGATGTGTGGGCGGCTTCGTGTACTGTGACGACCTGCTGGAGCAGGAGCTGTGC 1695
Qy 1756 GGTGTCTAGCGGCGCCAAACCCCGGGCGGGCTTACTGTGAGCAGATGCCCTATCCGT 1815
Db 1696 GGTGTCTAGTGGCAGGAGCTCCAGAGCAGGCTCTATTTGAGCAGATGCCACACCT 1755
Qy 1816 GCTATGTGGACGACGCTGTCTGCTGTGAGCGGCTGCTGCGGCTCTTCTTCTGAGCG 1875
Db 1756 GCTAGTGGACGATGTGTCTGCGGTGCTGAGCGGCTCTGCTGCTCTGTTTCTGATC 1815
Qy 1876 TGGCTGATCTTACTTCTGATGACATGACAGTGAAGCGCTGTGTCGGAGAGAGAGCG 1935
Db 1816 TGGCTGATCTTACTCGGTGCGCTCACTGTGAAGCGCTGTGTCGGAGAGAGAGAC 1875
Qy 1936 GGTGCGGACACCATGCGGCCCATGCGGCTCAGCGCGGCTGCTGCGCTAGGCTGT 1995
Db 1876 GGTGCGGAGAACCATGCTGCGATGGGGTTGAGCGGCTGTGCTGCTGCTTGGTGT 1935
Qy 1996 TCCTCAGCTGCTCGGGCTTCTGCTCAGCGCGCTGCTGCTGCTGCTGCTCAAGC 2055
Db 1936 TCCTCAGCTGCTGGGACCTTCTTGTGTCAGCGCTGCAATTTGCTGCTTAAAGC 1995
Qy 2056 TGGGGGACATCTCCCTACAGCACCGGCGGTGCTTCTGCTTGTGCGAGCTTGG 2115
Db 1996 TAGGGAACATCTCTTACAGCACCGGCTGTGCTGCTTCTTCTTGGCGGCTTGG 2055
Qy 2116 CGGTGGCCACGCTGACCCAGAGCTTCTGCTCAGCGGCTTCTTCTCCCGCGCA 2169
Db 2056 CGGTGGCCACGCTGCTCAGAGTCTTCTGCTCAGTGCCTTCTTCTCCCGGCGA 2109

RESULT 5

AK051920 LOCUS 4783 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length enriched library, clone:D330019D04 product:ATP-binding cassette, sub-family A (ABC1), member 1, full insert sequence.
ACCESSION AK051920.1 GI:26342297
VERSION AK051920.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Tanaka, T., Matsuda, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipette sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE
Functional annotation of a full-length mouse cDNA collection

REFERENCE
Nature 409, 685-690 (2001)

AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

REFERENCE
Nature 420, 563-573 (2002)

AUTHORS
6 (bases 1 to 4783)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Yamanatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission

REFERENCE
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
Location/Qualifiers
1. .4783

FEATURES
source

CDS

ORIGIN	Query Match Best Local Similarity Matches 1937; Conservative	13.3%; 56.5%; 0;	Score 904.4; Pred. No. 3.7e-175; Miematches 1371;	DB 3; Indels 120; Gaps 9;	Length 4783;
Qy	799	TGCGCAGCCTGTGTGGGCACTTCTGCGAGAGACCCGAGGAGCAGCGGCCCCC	858		
Db	1163	TGCGGAGGAGGTGATGTTTCTGACCAACGTGAACAGCTCCAGCTCCTCCACCCAGATCT	1222		
Qy	859	TGAGTTGCTGTACAGAGCCCTCTGCAGTGTACAGGGACCTACGACGACAGTGGGCCCT	918		
Db	1223	ACCAGCAGTGTCCCGCATCGTGTGTGTCAACCCAGAGGTTGGGGCTTGAAGATCAAGT	1282		
Qy	919	CCCTCAACTGGTACGAGGCTAGTGAACCTGATGAGCTGTGTGGG-----GCAGG	966		
Db	1283	CCCTCACTGTACGAGGATAACAACTACAAAGCCCTCTTTGAGGGAATAACACCGAGG	1342		
Qy	967	AGCCAGATCCGCCCTGCCAGACAGCAGCCTGAGCCCGCCTGCTCGAGCTGATTTGGAG	1026		
Db	1343	AAGACGTGGACACCTTCTATGACAAATTTCAAACTCTTATTGCAATGATTTGATGAAGA	1402		
Qy	1027	CCCTGGACGCCACCGCTGTCCGCTGTCTGTGGAGACCCCTGAAGCCCTCTGATCTCTCG	1086		
Db	1403	ACTTGGAGTCTAGTCTCTCTTCTCGAATTAATTTGGAAGCACTCAAGCCACTTCTTGTGTG	1462		
Qy	1087	GGAAGCTACTCTTTTGACCAAGATACACCTTTTACC CGGAAGCTCATGGCCAGGTGAACC	1146		
Db	1463	GAAGAATTCCTATACCTTACACACACCACTACAAGGCAAGGTCTGCGCTGAGGTGAACA	1522		
Qy	1147	GGACCTTCGAGGAGCTCACCTCTGAGGATGCTCGGAGGTGTGGGAGATCTCTGGAC	1206		
Db	1523	AGACCTTTACGAGTGTGGCTGTGTCATGACCTGGAGGCGCATGTGGGAAGAACTCAGCC	1582		
Qy	1207	CCGCGATCTTTCACCTTCATGAACGACAGTTTCCAAATGTGGCCATGTGCGACGCGCTCCT--	1264		
Db	1583	CCCAAATTTGACCTTCATGGAGAACGCAAGAGATGACCTTGTCCGACGCTGTAG	1642		
Qy	1265	-----GCAGATGCGAGTGAAGGAAGGAGGAGCCCA	1296		
Db	1643	ACAGCAGGCAATGACCAATTTTGGGAAACAGAAAGTTGGATGATTAGATTGACTGCCCC	1702		
Qy	1297	GACCTGGAGCCGGGACCAATGAGGCCCTCGATCTCTTCTGGACCTCTGGGAGCGGTG	1356		
Db	1703	AAGACATCATGGCGTTTCTGCCAAGAACCCAGAAAGTGTCTCACTCCCCAAATGGCTCTG	1762		
Qy	1357	GCTACAGCTGGCAGGACGACACACGCTGATGTGGGACACCTGTGTGGGACGCTGGGCCGAG	1416		
Db	1763	TGTATACCTCGAGAGAAGCTTTCAATGAGAACCAACAGGCAATCCAGACGATATCTCGAT	1822		
Qy	1417	TGACGAGTGCCTGCTCTTGGACAAGCTGGAGCGGCACCCCTCAGAGGAGCCCTGTGT	1476		
Db	1823	TCATGAGTGTGTCAACCTGAAACAGCTGGAAACCCATTCCGACAGAAAGTCAGGCTCATCA	1882		
Qy	1477	CGCGGCCCTTGCACCTGCTCGCGGAACATCGAATTTCTGGCGCGCGTCTCTTCTTGGGAC	1536		
Db	1883	ACAAGTCCATGGAGCTCTGGAACGAGGGAAGTTCTGGGCTGGCATCGTGTTCACAGGC-	1941		
Qy	1537	CTGAGGACTCTTACAGACCCACAGAGCAACCAACCCAGACCTTGGGCCCGCGGACGCTGC	1596		
Db	1942	-----ATCACTCCAGATGTGTGGAGTGTGCCCATCATGTAA	1978		
Qy	1597	GCATCAAAATCCGATGACATTTGACGTGTGTACGAGGACCAATAAGATCAGGACAGGT	1656		

1979 AGTAAAGATCCGGATGGACATTGACAACTGGAGAGAACTAATAAGATCAAGGATGGT 2038
1657 TTGTGGACCTCGCCAGCCGCGGACCCCTGACCGACTCGCTACGTTGGGGGGCT 1716
2039 ACTGGGACCTCGCTCGTGGGCTGACCCCTTTTGAAGATATGGCTATGTTGGGGGGCT 2098
1717 TCGTGTACTGCAAGACCTGTGTGGAGCGTGCAGCCGCTCGCGTGTCTACGCGCGCAACC 1776
2099 TCGCCTACTTGCAGGATGTGTGGAAACAGGCCATCATCAGAGTGTGACGGGATCTGAGA 2158
1777 CCCGGCCGCGCTCTACCTGACGAGATGCCCCCTATCCGTGCTATGTGACAGACGTGTCC 1836
2159 AGAAAAACGGGTGTCTACGTGCAACAGATGCCCTTACCCCTGTTATGTTGATGACATTTTTC 2218
1837 TCGTGTGTGTGACCGCGCTCGCTCGCGCTCTTCTGACGCTGCGCTGGATCTACTCCGTA 1896
2219 TCGGGGTATGAGCCCGGTCAATGCCCTCTTCTGACTGTAGCTGGATCTACTCTGTGC 2278
1897 CACTGACAGTGAAGCCCGTGTGTGGGAGAGAGACGCGCTGCGGACACCATGCGCG 1956
2279 CTGTATCATCAAGAGCATTTGTATGAGAGAGGCTCGGCTGAAGAGACCATGCCGA 2338
1957 CCATGGGCTCAGCCGCGGGTGTCTGTGGCTAGGCTGTGTTCTCAGCTGCTTGGGCCCT 2016
2339 TCATGGGTCTGGACAAATGGCATCTCTGTGTTAGCTGTGTTGTTAGCAGCCTCATCCCTC 2398
2017 TCCTGCTCAGCGCCGCTGTGTGTTCTTGGAGCCCTTGGCGTGGCCACGCTGACCCAGA 2136
2459 GTGACCCAGCGTGTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2518
2137 GCTTCTCTCTCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2196
2519 GCTTCTCTCTTAGCACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2578
2197 CCTACTTCTCTCTACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2256
2579 TCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2638
2257 CGGGTGGCGGCTGGCCGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2316
2639 TCTCCATCAAGATCTTTGCTAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2698
2317 GCCTGGCTCTCTGGAGAGAGCGGCGAGGCGCGCAGTGGACAA ---CGTGGCACCC 2373
2699 ATTTGGCCCTTTTCGAGGAGCAAGGTATCGGGGTCCAAATGGGACAACTCTCTTTGAGAGCC 2758
2374 GGCCTACGGCAGAGTCTTTACGCTTGGCCCAAGTCTCTGGCTCTCTCTCTCTCTCTCTCT 2433
2759 CGGTGGAGGAGAGCGCTTCAATCTCACTCACTGAGTGTCTATGCTCTTTGACACCT 2818
2434 CGCTCTACGCGCTCGCCACCTGTGTACCTGGAAGTGTGTGCGCCAGGCGCAGTACGGGATCC 2493
2819 TTCTCTATGGCGTGATGACATGGTATCATCGAAGCCGTCTTTCCAGGACAGTATGGAATTC 2878
2494 CTGAACCATGAAATTTCTTTTTCGGAGGAGCTACTGTGTGGGACCTTGGCCCCCCCCAAGA 2553
2879 CCAGGCCCTGTATTTCTCTGTACCAAGTCACTACTGTGTTGGTGGAGAAATTTGATGAGA 2938
2554 GTCCAGGCCCTTGGCCACCCCGCTG ---GACCCAAAGTGTCTGGTAGAGAGGACCGC 2610
2939 AGAGCCACCTTGGTTTCAGCCAGAGGAGGTGTGAGAAATCTGATGGAAGAGAAACCCA 2998
2611 CCGGCTGTAGTCTCTGGGCTATCTCTGCTGACCTGGAGAGCGCTTTCTCTGGAAGCCGC 2670
2999 CTCACTGTAGGCTGGGGGTGTCATTTAGAACCTGTGTGAAGTTTACGAGATGGCATGA 3058
2671 AGCCAGCCCTGCGGGGCTCAGCTGTGACTTCTACAGGGCCACATCACCCTTCTCTGG 2730

3059 AGTGTGCTGTGGATGGCTTGGCGCTCAACTTTTACGAAGGCCAGATTACTCTCTTCTCTGG 3118
2731 GCCACAAACGGGGCCGGCAAGACCAACCCCTGTCTCATCTTGTAGTGGCTCTTCCCAACCCA 2790
3119 GCCAAATGGAGCAGGGAAGACCAACCATGTCAATAGTACTGACTGGGCTGTTTCCCCCAA 3178
2791 GTGTGGCTCTGCGCTTCACTCTGGGCCACGAGCTCGCTCCAGCATGGCGGCATCCCGC 2850
3179 CTTCTGGCAACCGCTTACATCTGGGAAGGACATTCGCTCGGAGATGAGCTCATCTCCGC 3238
2851 CCCACTGGGCGTCTGTCTCTCAGTAAACGCTGCTGTTTGAACATGCTGACCGTGGAGC 2910
3239 AGAACTGGGAGTCTGTCCCGACATAATGTCTGTTTGAATGCTGACTGTGCGAAGAGC 3298
2911 ACCTGTGTTCTATGGCGGCTGAAGGCTCTGAGTCCGCTGTAGTGGGCCCGCGAGCAG 2970
3299 ACATCTGGTTCTATGCGCGCTTAAAGGGGCTCTCAGAGAAACGACGTGAAGCAGAGATGG 3358
2971 ACCGTCTCTGCAGGATGTGGGGCTGTCTC ---CAAGCAGAGTGTGCAGACTCGCCACC 3027
3359 AGCAGATGGCCCTGGATGTGTGCTTACCCCGAGCAGCTGAAAAGCAAAAGAGTACG 3418
3028 TCTCTGTGGATGCAACGGAAGCTGTCCGTGGCAATTCCTTTGTGGCGGCTCTCCCAAG 3087
3419 TCTCAGTGGGATGACAGAGAAAGCTGTCTGTGGCTTGGCTTGGTGGATCCAAG 3478
3088 TTGTTATCTCGACGAGCTACGCTGGCTGGATCTGTCTTCCCGCGCGGTATTTGGG 3147
3479 TTGTCAATCTGACGAGGCCACGCGGGGTGGACCGCTACTCTCGCAGGGGAATATGG 3538
3148 AGCTGTCTCTCAAAATACGGAAGGTGCACGCTGATCTCTCCACCAACCACTGGATG 3207
3539 AACTCTCTGCTAAATACGCGCAGAGCCGCAATTTTGTCTACACACACATGGAGC 3598
3208 AGGAGAGCTCTCGGAGACCGTGTGGCTGTGGTGGAGTGGCGCTTGTCTCTGTG 3267
3599 AAGCTGACATCTTTGGGAGACAAATTTGCCATCAITTTCCATGGGAAGCTGTGTGTGG 3658
3268 GCTCCACATCTTCTCGCGCTCACCTGGGCTCGGCTACTACTGTGACGCTGGTGAAG 3327
3659 GCTCTCTCTCTTTTGAATAACAGTTGGGAAACGGGTACTATCTGACCCCTGGTTAAG 3718
3328 CCCGCTCTCTCTGACCAACCAAT ---G 3351
3719 AAGATGTGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3778
3352 AGAAGGTGACATGACATGGAGGGCAGTGTGGACAACAGCAGGAAAGAAAGATGACA 3411
3779 AAAAGGAGGACAGTGTCTCTCAGAGCAGTCTGATGTCTGGCTGGGACGACCATGAAA 3838
3412 GCCAGGGCAGCAGAGTGGGACCTCTCAGCTGTGGCCCTGGTACAGCACTGGGTGCCCG 3471
3839 GTGACACGCTGACCATCGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3898
3472 GGGCACCGCTGGTGGAGGAGCTGCACACGAGCTGTGTCTGTCTGTCTGTCTGTCTGTCT 3531
3899 AAGCCCGCTGGTGGAGGACATTTGGGACAGAGCTGACCTATGTCTGTCTGTCTGTCTGT 3958
3532 CCCATGACGGCAGCTTCCGCAACTCTTCCGAGAGCTAGACACGCGCTGGCGGAGTGA 3591
3959 CGAAGGAGGAGCCCTTTGTGGAACTCTTCCATGAGATTGATGACCCGCTCTCAGACCTGG 4018
3592 GGCTCACTGTGCTACGGGATCTCCGACACCGACCTCTGAGGAGATCTTCTGAGGTGGTGG 3651
4019 GCATCTCAGTTATGGCATCTCTCGAGACCACTCTGGAAGAAATATCTCTAAAGTGGCTG 4078
3652 AGGAGTGTGTCTGCGACACAGATATGGAGATGGAGTGGAGTGGCGGACGACCTATGACAG 3711
4079 AAGAGAGCGGGTGGATGTCTGAGACCTCAGATGTTATTTTCCAGAGAGAGCAACAGAC 4138
3712 GCATGTCTGAGAGCTTAACCTCTGGCTTAAGATGCGGCCACAGAGAGACAGC ---GC 3768
4139 GGGCTTTCGGGGAACAGCAGAGCTGTCTGCACCCATTTACGGAAGATGATGCTGTGTATC 4198

QY 3769 TGGAGAACGGGGAACACAGCTGGGTACGCCAGAGACTGACCCAGAGAGTACAGAGGGCTCTGGGCCAGACG 3828
Db 4199 CCAATGACTCTGACATAGACCCAGAAATCCAGGGAGACCGACCTGCTCAGTGGGATGGACG 4258
QY 3829 CCGTGGG-----CCGGGTACAGGCTGGGCACTGACCCGCCAGCAGCTCCAGGCGCTGC 3882
Db 4259 GCAAGGCTCTTACAGCTGAGAGGGCTGGAACTCACCCAGCAACAGTTTGTGGCCCTTT 4318
QY 3883 TTCTCAAGCGCTTTCTGCTTTGCCGCGCAGCCGCGGCTGTTTGGCCAGATCGTGC 3942
Db 4319 TGTGGAAGAGGCTGCTGATTCGACAGCGAGCCGGAAGGGTTCTTTGCTCAGATTGTCC 4378
QY 3943 TGCCTGCTCTTTTGTGGGCTGGCCCTGCTGTTTACGCTCATGCTGCTCTTGGGG 4002
Db 4379 TGCAGCTGTCTTTTGTGATGCTGCTGCTTTCAGCTGATTTGTGCCACCCCTTTGGCA 4438
QY 4003 ACTACCCGGCTCTGGGCTCAGTCCACCACTGATACGGTGTCTCAGGTGCTCTTCAGTG 4062
Db 4439 AGTACCCCAAGCTTGAACCTTCAGCCCTGGATGTATATAGACAGTATACATTTGTCA 4498
QY 4063 AGGACGCCGCCAGGGAACCTTGGAGCTGCCGCTGCTCGAGGGCTGCTGCAGGAGCGAG 4122
Db 4499 ATGATGCTCCGAGGACATGGCACCCAGGAACTCTCTGATGCTCTGACCAAGATCCAG 4558
QY 4123 GACTGGAG 4130
Db 4559 GCTTTGGG 4566

RESULT 6

BQ959013

LOCUS AGENCOURT 8958795 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6460431
DEFINITION 5', mRNA sequence.

ACCESSION BQ959013

VERSION BQ959013.1

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 933)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM2647 row: b column: 16

High quality sequence stop: 677.

Location/Qualifiers

1. 933

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6460431"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_101"

/note="Organ: lung; Vector: pOTB7; Site: 1: EcoRI; Site 2:

XhoI; CDNA made by oligo-dr priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library "

ORIGIN	Query Match	Best Local Similarity	Score	828.8;	DB 5;	Length	933;			
	Matches	884;	Conservative	0;	Mismatches	33;	Indels	6;	Gaps	3;
QY	2982	CAGGATGTGGGGTGTCTCTCAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATG	3041							
Db	1	CAGGATGTGGGGTGTCTCTCAAGCAGAGTGTGCAGACTCGCCACCTCTCTGTGGGATG	60							
QY	3042	CAACGGAAGCTGTCCTGGGCCATTGCTTTGGGGCGCTCCCAAGTTGTTATCTCTGGAC	3101							
Db	61	CAACGGAAGCTGTCCTGGGCCATTGCTTTGGGGCGCTCCCAAGTTGTTATCTCTGGAC	120							
QY	3102	GAGCCTACGGCTGGCGTGGATCTGTTCCCGCGCGGTATTTGGGAGTGTCTCTCAAA	3161							
Db	121	GAGCCTACGGCTGGCGTGGATCTGTTCCCGCGCGGTATTTGGGAGTGTCTCTCAAA	180							
QY	3162	TACCGAAGGTGCGACGCTGATCTCTCCACCCACACCTGATGAGGAGAGCTGCTG	3221							
Db	181	TACCGAAGGTGCGACGCTGATCTCTCCACCCACACCTGATGAGGAGAGCTGCTG	240							
QY	3222	GGAGACCGTGTGCTGTGTGGCAGGTGGCGCTTGTGCTGTGTGTGTGTGTGTGTGT	3281							
Db	241	GGAGACCGTGTGCTGTGTGGCAGGTGGCGCTTGTGCTGTGTGTGTGTGTGTGTGT	300							
QY	3282	CTCGCCCGTCACTGGGCTCGGCTACTACTGACGCTGGTGAAGGCCCGCTGCCCTG	3341							
Db	301	CTCGCCCGTCACTGGGCTCGGCTACTACTGACGCTGGTGAAGGCCCGCTGCCCTG	360							
QY	3342	ACCACCAATGAGAAGCTGACACTGATGGAGGGCAGTGTGACACCCAGGAGGAAAG	3401							
Db	361	ACCACCAATGAGAAGCTGACACTGATGGAGGGCAGTGTGACACCCAGGAGGAAAG	420							
QY	3402	AAGAAATGGCAGCCAGGCGCAGCAGTGGCAGTCTCCTCAGCTGTGGCCCTGGTACAGC	3461							
Db	421	AAGAAATGGCAGCCAGGCGCAGCAGTGGCAGTCTCCTCAGCTGTGGCCCTGGTACAGC	480							
QY	3462	TGGGTGCCGGGGCAGCGGTGGTGGAGGAGCTGCCACACGAGCTGGTGTGTGTGTGT	3521							
Db	481	TGGGTGCCGGGGCAGCGGTGGTGGAGGAGCTGCCACACGAGCTGGTGTGTGTGTGT	540							
QY	3522	TACACGGGTGCCCATGACGCGCAGCTTCCGACACTCTTCCGAGAGCTAGACACGCGCTG	3581							
Db	541	TACACGGGTGCCCATGACGCGCAGCTTCCGACACTCTTCCGAGAGCTAGACACGCGCTG	600							
QY	3582	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACGAGCTCGAGGAGATCTTCTG	3641							
Db	601	GCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACACGAGCTCGAGGAGATCTTCTG	660							
QY	3642	AAGGTGGTGGAGAGTGTGTCGGACACAGATATGAGGATGGCAGCTCGGGGAGCAGC	3701							
Db	661	AAGGTGGTGGAGAGTGTGTCGGACACAGATATGAGGATGGCAGCTCGGGGAGCAGC	720							
QY	3702	CTATGACACAGGATTTGCTGGCTTAGAGTTAACTTGGGCTCAAGTTCGCGCAGCAGAG	3761							
Db	721	CTATGACACAGGATTTGCTGGCTTAGAGTTAACTTGGGCTCAAGTTCGCGCAGCAGAG	780							
QY	3762	ACAGC-GCTGGAGAACGGGG--AACAGCTGGGTTCAGCCGCCAGAGACTGACCGAGGCTCT	3818							
Db	781	ACAGCGCTGGAGAACGGGGGAAACCCAGCTGGGTTCAGCCGCCAGAGACTGACCGAGGCTCT	840							
QY	3819	GGGCCAGACCGCGCTGGG---CCGGGTACAGGGCTGGGCACTGACCCGCCAGAGCTCAG	3875							
Db	841	GGGCCAAACCCCGCGGGGCCCGGGGTACAGAGGGCTGGGCACTGCAACCCCTTCACTTCG	900							
QY	3876	GCCTGCTTCTCAAGCGCTTTCT 3898								
Db	901	GCCTGCTTCTCAAGCGCTTTCT 923								

RESULT 7

CR627391
LOCUS CR627391 3759 bp mRNA linear HTC 03-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFp781N1972 (from clone DKFp781N1972).
ACCESSION CR627391
KEYWORDS HTC.
SOURCE HTCC.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3759)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeinger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFp781N1972) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFp781N1972
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

1..3759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFp781N1972"
/tissue_type="retina"
/clone_lib="781 (synonym: hicc4). Vector pSport1_sfi; host
DH10B; sites sfIIA + sfIIB"
/dev_stage="adult"
/note="ATP-binding cassette, sub-family A member 4,
differentially spliced"

gene

1..3759
/gene="DKFp781N1972"
141..3338
/gene="DKFp781N1972"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10486.1"
/db_xref="GI:50949876"

/translation="MDVVVHHVPEAKLVEICQELIFLLPNKFKHAYASLIFRELEE
TLADLGLSSFGISDTPLBEEIFLKVTEDSGLPAGGAQKRENVNPRHPCGLPREKA
GOTQDSNVCSGPAAPAHPEGPPPEPCGPQNTGTLVLQHVQALLVKRFOHTLR
SHKDFLAQIVLPATFVFLALMSIVIPFGYPAITLHPWIYQQYTFPSMDEPGSEQ
FTVLADVLNPGFNGCLKEGMWPEYPCGNSPTKMTSPVSPNITQULPQKOKWQVNP
SPSCRSREKLTMLPECEGAGLPPQRTORSTELQDLTDRIISDFLVKTYPALI
RSSLSKTFVNEVYRGYISGIGLPPVITGEALVGLSLGRIMNVSGSGPITREASK
BIPDLKHELTEDNIKFNWKNHVALVIRASLPRKDRSPSEYIGITVIS
QPLNLTKEQLSEITVLTVSDVAIVLCVIFSMFVPSFVLVLIQERNVSKHLQFIS
GVSPITYWTFNLFMDIMNYSVAGLVGIFGQKXATSPENLPALVALLLGYAV
IPMTYPASFLDVSTVALSCANLFIQINSALITFLEFPENRNLRLRNVLKRL
LIVPFHFLGRGLDIALUSQAVTVYARFGEHSANPFHDLIGNLPAMVVEGVYF
LLTLVQHFGLLSQIAHPTKEPIVDEDDVAERQRIITGNKTDILRLHELTKIYP
TSSGPVNLRCVGRPGECFGLLVGSGAKTTFKMLTDTVTSGDATVAGKSLTN
ISEVHONGYCPQDAIDELLTGREHLVYARLGRVPAEEIEKVANNVSKLGLTVYA
DCLAGTSGGNKRKLSTALIALIGCPPLVLLDEPTGMDPQARRMLNVVSIIREGRA
VLTLSHSECEALCTRLAIWVGAFCRQVGTIQLHLSKFGDGIYIVTMKIPKDDLLP
LDNVPQFQFENFQSVGRHYNNMLQFVSSSLARIFQLLSHKDLSLLIEYSVTQ
TTLDQVFNFAKQQTESHDLPLHPRAGASRAQAD"

ORIGIN

Query Match 12.1%; Score 820.2; DB 3; Length 3759;
Best Local Similarity 55.7%; Pred. No. 7.8e-158;
Matches 1747; Conservative 0; Mismatches 1336; Indels 54; Gaps 7;

3437 TCAGCTGCTGGCCCTGGTACACACTGGGTGCCCGGGCACCGCTGGTGGAGAGCTGCC 3496
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 193
3497 ACACAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3556
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TCAAGAACTTATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 253
3557 CTTCCGAGAGCTAGACACGCGGCTGGCGAGCTGAGGCTGCTCTGCTGCTGCTGCTG 3616
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TTTTCAGAGAGCTGGAGGAGAGCTGGCGGACCTTGGTCTTCAGCACTTTTGGAAAT 313
3617 CACCAGCCTCGAGGAGATCTTCTGAAAGTGGTGGAGGAGTGTGCTGGGACACAGAT 3676
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db CACTCCCTGGAAGAGATTTTCTGAAGTCAOGGAGGATCTGATTGATCAGGACCTCT 373
3677 GGAGATGGC-----AGCTGGGGGACGCACTATGCAAGGCACTGCTGGCTAGAGCT 3730
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TCGGGTGGCGCTCAGCAGAAAGAAAGAGAGCTCAACCCCGACACCTCTGCTGGGTCC 433
3731 AACCTCGGGCTCAAGATGCCGCCACAGGAGACAGCGCTGGAGAACGGGAACAGCTGG 3790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db CAGAGAGAGGCTGGACAGACACCCAGGACTTCAATGTCTGCTCCCAAGGGGCCCGGC 493
3791 GTCAGCCCCAGAGACTGACACAGGGCTCTGGGCCAGACCGCGTGGCGGTACAGGGCTG 3850
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db TGCTCACCAGAGGGCCAGCTCCCGCAGAGCCAGAGTGCCAGGCCGAGCTCAACAC 553
3851 GG-----CACTGACCCGCCAGAGCTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTGGC 3904
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db GGGGACACAGCTGGTCTCCAGCATGTGCAGCGCGTGTGGTCAAGAGATTCCAACACAC 613
3905 CGCGCAGCGCGCGCGCGCTTTCGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3964
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db CATCGCAGGCCACAAGGACTTCTCGCGCAGATGCTGCTGCTGCTGCTGCTGCTGCTG 673
3965 GGCCTCTGTTGACGCTCATCGTGCCTCTTTTGGGCACTACCGGCTCTGCGGCTCAG 4024
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db GGCTCTGATGCTTCTATTGTTATCCCTCTTTTGGCGAATACCCGCTTTGACCTTCA 733
4025 TCCACCATGTACGGTGTCTAGGTGTCTTCTTCAAGTGAAGAGCGCCAGGGGACCTGG 4084
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db CCCCTGGATATATGGCGAGCAGTACACCTTCTTCAAGCATGGATGAACAGGCGAGCA 793
4085 ACCTGCCCGGTGCTCGAGGCGCTGCTGAGGAGCAGGCTGGAGGAGCCCGCAGTGCA 4144
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db GTTCAGGTACTTGCAGACGCTCTCTGAAATAAGCCAGGCTTGGCAACCGCTGGCTGAA 853
4145 GCA-----TAGCTCCCAAGGTTCTCGGCACACAGAGT 4177
854 GGAAGGTGGCTCCGGAGTACCCCTGTGGCAACTCAACACCTCGAAGACTCCTTCTGT 913
4178 TCCTGCTGAATGGCAAGGTCTTGGCCAGTGGGCAACTGGAGCCCGCAGAGTCTCCATCC 4237
Db GTCCCCAACAATCAACCCAGCTGTTCCAGAAGCAGAAAATGGACACAGGTCAACCCCT 973
4238 AGCTGCCAGGTAGCCAGCCCGGTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCGAGC 4297
Db ATCTGCGAGGTGACAGCACCAGGAGAGCTCACATGTGCCAGAGTGCCCGCAGGGTGC 1033
4298 TGGTGGTCCCCCTCCGCCCGCAGGAGTGAACCGGCTCTGGGGAAGTGGTTTCAGAACTGAC 4357
1034 CGGGGGCTCCCGCCCCCCCCCAGAGAAACACAGCGCAGCAGCGAAAATTCACAGACCTGAC 1093
4358 AGGCGGAACTGTCTGATCTTCTGGTCAAGACTTACCCCGCGCTGGTGGCGCAGGGCT 4417
Db GGACAGGAACTCTCCGACTTCTTGGTAAAAACGATATCCTGCTCTTATAAGAACGAGCTT 1153
4418 GAAGACTAAGAGTGGGTGAATGAGTTCAGTACGAGGCTTCTCGCTGGGGGGCCGAGA 4477
Db AAAGAGCAAAATTCGGGTCAATGAACAGAGGTATGGAGAAATTTTCATTGGAGGAAGCT 1213
4478 CCCAGGCTGCC---CTCGGGCCAAGAGTTGGGCGCTCAGTGGAGGAGTTGTGGGGCGCT 4534

[illegible][illegible]

RESULT 8
AK052916
LOCUS
DEFINITI

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK052916
AK052916.1 GI:26095451
HTC: CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS

2 Carninci, P., Shibata, Y., Hayateu, N., Sugihara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
11076861

TITLE

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
11076861

TITLE

6 (bases 1 to 3605)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Haseizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koye, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp).
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>

FEATURES
source

URL: <http://fantom.gsc.riken.jp/>.
Location/Qualifiers
1..3605

/organism="Mus musculus"
/mol_type="mRNA"
/strains="CS7BL/6J"
/db_xref="PANTOM_DB:D830040G08"
/db_xref="taxon:10090"
/clone="D830040G08"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
misc_feature
1..3605
/note="ATP-binding cassette, sub-family A (ABC1), member 4
(MCD|MGI:109424, GB|NM_007378, evidence: BLASTN, 99%,
match=3552)"

ORIGIN

Query Match 11.8%; Score 805.8; DB 3; Length 3605;
Best Local Similarity 55.8%; Pred. No. 7.2e-155;
Matches 1762; Conservative 0; Mismatches 1327; Indels 70; Gaps 9;
QY 3439 AGCTCTGGCCCTGTGTACAGCACTGGTGGCCGGGACAGGCTGTGTGGAGAGTCCAC 3498
Db 65 AGCTGATGGATTGGTATATACCACTATGCCAGAGGCAAGCTGGTGGAAATGCAATGGTC 124
QY 3499 ACGAGCTGGTGTGTGTGCTTACAGGGTGGCCATGACGGCAGCTTGGCCACTCT 3558
Db 125 AAGAACTTATCTCTCTCCCTCCNAACAAAGATTTCAGAGAGAGCATATGCCAGCTTT 184
QY 3559 TCCGAGAGCTAGACACACGGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCGACA 3618
Db 185 TCCGAGAGCTGGAGGAGACTCTGGCTACCTGGGCTCAGCAGCTTGGAAATTCGACA 244
QY 3619 CCAGCTCGAGAGATCTTCTGAAGTGGTGGAGGAGTGTGTGCGGACACAGATATGG 3678
Db 245 CTCCCTCGAAGAGATTTTCTGAAAGTCAAGAGGATGCTGGAGCGGCTCTATGTTG 304
QY 3679 AGGATGGCAGCTGGGGCAGCACCTATGCAAGGCAATGCTGGCTAGACGTAACCTGC 3738
Db 305 TAGTGGCGCTCAGCAGAAAGAGAAACAAAGCGGCTCTCGGACCCCTGTTGGCTGCCA 364
QY 3739 GGCTCAAGATCGCCACAGGAGACAGCGCTGGAGAACGGGAAACAG-----CTGGGTC 3793
Db 365 CTGAGAGCTCAGGAGTATGCCAGGCCCACTACCTGTTCCAGGACAGTGGATC 424
QY 3794 AGCCCCAGAGACTGACAGGCTCTGGGCCACAGCCGCTGGGCCGGGTACAGGGTGGG- 3852
Db 425 CTCCCAAGGGCCAACTTCCCGCAGAGCCAGAACGCCGGTGTCCATTCAACACAGGTG 484
QY 3853 ---CACTGACCGCCAGCAGCTCCAGGCCCTGCTTCTCAAGGCTTCTGCTTGGCCGCC 3909
Db 485 CTGGCTGATTCTTCAACATGTGAGGCGCTGTAGTCAAGCGATTCCATCACACCATCC 544
QY 3910 GCAGCCGGCGCGCTGTTCGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3969
Db 545 GCAGCCGAGAGACTTGTGGCTCAGATGTTCTCCCTGCCACTTCTGTTCTTGGCTC 604
QY 3970 TCGTGTTCAGGCTCATGTGCTCTTTCGGGCACTACCCGGCTCTGGGCTCAGTCCCA 4029
Db 605 TGATGCTTCAATCATTTGTGATCCGTTGGTGAATTTCCAGCTTGTACTCTTCATCCCT 664
QY 4030 CCATGTACGCTCAGTGTCTTCTTCAAGTGGAGCGCCCGGAGCCCTGGACCTG 4089
Db 665 GGATGTATGGGATCAATATATACCTTCTTCAGATGGAAGAACCCCAACATGAACACCTTG 724
QY 4090 CCCGGCTCTCGAGCGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4145
Db 725 AAGTACTGGCAGATGCTCTTCTGAAAGCCAGGCTTGGCNAACCGTTGTCTTAAGAGAG 784
QY 4146 CATAGTCCCAAGGTTCTCGGCAACAGAGTGTCTGCTG----- 4185
Db 785 AGTGGCTTCCGAGTACCCTATGATGCAACCTCTGGAAGAGCGCCCTCCGTGTC 844

Qy 4186 ---AAGTGGCCAAAGGTTCTTGGCCAGTGGCACTGGACCCAGAGTCTCCATCCCAAGCCT 4242
Db 845 CGAACATACACCCACTGTTCAGAAAGCAGAAATGGACAGACGCCACCCCTCTCCCTCT 904
Qy 4243 GCCAGTGTAGCAGCCCGGTGCCCGCGCTGCTGCCGAGTGGCCGAGTGGCTGCAGCTGGTG 4302
Db 905 GCAAGTGAGCACCAGAGAGAGAGCTCACCATGTGTGCCGAGTGGCCGAGGGTCTGGAG 964
Qy 4303 GTCCCCCTCCGCCCCAGCAGTGCACCGCTCTGGGGAAGTGGTTCAGAACTCTGACAGGCC 4362
Db 965 GGCTCCCAACCCACAGAGAGACACAGCGCAGCACTGAAGTCTTACAGACTCTCAGAAC 1024
Qy 4363 GGAACCTGTCTGACTCTCTGTGTGAAGACTTACCGCGCTGGTGGCGCAGGGCTGAAGA 4422
Db 1025 GGAACATCTCCGACTACTTGGTAAAAACGTACCCCTGTCTCTATAAGAGAGCAGCTTAAAGA 1084
Qy 4423 CTAGAAGTGGGTGAATGAGGTACAGGTACGAGGCTTCTCGCTGGGGGCGGAGACCCAG 4482
Db 1085 GCAAAATCTGGGTCAATGAACAGAGGTATGGAGAAATTTCCATCGGAGGAAATCTCCCTG 1144
Qy 4483 GCCTGCC---CTCGGGCCAAAGAGTTGGGCGCTCAGTGGAGAGTGTGTGGCGCTGCTGA 4539
Db 1145 CTATCCCACTCAGCGGGGAGCACTTGTGGCTTTCTAAGTGGCTTGGCCAGATGATGA 1204
Qy 4540 GTCCCTGCTGGCGGGGCCCTCGACCGTGTCTGTGAATAAACCCTCAGAGCTGGGCTCACA 4599
Db 1205 ACGTAGCGGGGGTCTGTGTACCAAGGAGGCTCCAAAGAAATGTTAGATTTCTCTAAAC 1264
Qy 4600 GCCTGGATGCTCAGGACAGTCTCAAGATCTGGTTCACAAACAAAGGCTGGCACTCCATGG 4659
Db 1265 ATCTTGAAACACAGACACAAATTAAGTATGGTTTAACAAACAGGGCTGGCAATGCCCTGG 1324
Qy 4660 TGGCTTTTGTCAACCGAGCCAGCAACCAATTCCTCGTGTCTCACTGCCCCAGGCCCGG 4719
Db 1325 TCAGCTTTCTGAATGTGCTCAACACGCCATCTTAAGGGCCAGCTGCCCGAGGACAGGG 1384
Qy 4720 CCGCCAGCCACAGACATACACACACTCAACACACCCCTTGAACCTCACCAGAGAGCAGC 4779
Db 1385 ACCGTGAGGATATGGAATCACTGTCTATCAGCCAGCCTCTGAACCTGACCAAGAGCAGC 1444
Qy 4780 TGTCTGAGGTGCACTGATGSCCTCTCGGTGGAGCTCTCGTCTCCATCTGTGTGTCT 4839
Db 1445 TCTCGGACATCAGTACTGACTACTTCCGTGGATGCGGTGGTGGCCATATGTGTGATTT 1504
Qy 4840 TTGCCATGTCTTTGTCCCGCCAGCTTCACTCTGTCTCTCAATGAGGAGCAGTCAACC 4899
Db 1505 TCGCCATGTCTTGTCTTCCAGCCAGCTTGTGCTTTACTTGATCCAGGAGGGGTGACGA 1564
Qy 4900 GAGCAAGCACTGCAGTCAATGGGGGCTGTCCCGACCCCTCTACTGGCTTGGCAACT 4959
Db 1565 AAGCTAAACATTTGGCAGTTATCAGTGGTGTGAGCCCGCCACCACTACTGGCTGACCAACT 1624
Qy 4960 TTCTCTGGGACATGTGTAACTTACTTTGTGGCAGCATGATCGTGTGTCTCATCTTTCTGG 5019
Db 1625 TCCCTGGGATATTAAGAAATTAATCTGTGTAGTGGGATTTGGTGGGCAATCTTCAATTG 1684
Qy 5020 CTTTCCAGCAGAGGGCATATGTGGCCCTGTGCCAACCCTGCTGTCTCTCTGTGTGTGTAC 5079
Db 1685 GATTTCAAAAGAAAGCCTACACATCCCGACAGCAACCTTCCGTCTCTGTCTCTGTCTCA 1744
Qy 5080 TACTGTATGGTGTGTGATCAGACCGCTCATGTATACCAGCCTCTCTTCTTCTCCGTGC 5139
Db 1745 TGTGTATGGATGGGCACTATTCCTATGATGTATCCAGCATCTCTCTGTGTGAAGTCC 1804
Qy 5140 CCAGCAGCCTATGTGTGTCTCACTGATCAATAAACCTCTTTATGGCATCAATGGAAGCA 5199
Db 1805 CCAGCAGCCTATGTGTGTCTTGTCTGTGTAACTTCTTCACTCGGATCAACAGCAGCG 1864
Qy 5200 TGGCCACTTTGTGCTTGTAGCTCTT---CTCTGATCAGAAAGCTCAGAGGTGAGCCGGA 5256
Db 1865 CCATCACCTCTCGTCTCTGGAATTAATTGAGAAATACCGGACGCTGCTCAGGTTCATGCCA 1924

Qy 5257 TCTTGAAACAGGTTCTTCTTATCTTCCCCACACTTCTCTTGGGCGGGGCTCATTTGACA 5316
Db 1925 TGTGAGAAAGTTGCTCATTTGCTTCCCCACACTTCTGCTTGGGAGGGGCTCATTTGACC 1984
Qy 5317 TGTGCGGAAACAGGCTGATGCTTGTAGCCCTTGGGAGACAGGCAAGTCTCCAGT 5376
Db 1985 TGGCACTGAGCAAGCTGTGACAGAGCTCTATGCCAGTTTGGTGAGAGTACTCTGCNA 2044
Qy 5377 CACCCCTGCGCTGGAGGTGGTGGCAAGAACCTCTTTGGCCATGGTGATACAGGGGCCCC 5436
Db 2045 ACCCATTTCCAGTGGGACCTGATTTGGGAAGAACCTGTTGCTATGGCAATAGAGGGGTG 2104
Qy 5437 TCTTCTTCTTTCACACTACTGCTGACAGCACCGAAGCCAACTCTCTGCCACAGCCAGGG 5496
Db 2105 TATATCTTCTTCTGACCCCTGCTCATCCAACCACTTTTTCCTCACCGGTGSAATCGCTG 2164
Qy 5497 TGAGGTCTCTGCCACTCTCTGGGAGAGGAGCAGAGGATGTAGCCCGTGAACGGGAGCGGG 5556
Db 2165 AGCTGTCTAGAGAGCCCGTTTTTGTATGAGATGACGATGTGGCTGAAGAAAGACAAAGAG 2224
Qy 5557 TGGTCCAAAGAGCCOACCCAGGGGGATGTGTGTGTGTGAGGAACTTTGACCAAGATATACC 5616
Db 2225 TTATGAGTGGGGGAAATAAAACCGACATCTTAAAGCTTAAACGAACTAAACCAAGTTTACT 2284
Qy 5617 GTGGCAGAGGATGCCAGCTGTTCACCGCTTGTGCCCTTGGGATTTCCCTCTGCTGAGTGT 5676
Db 2285 CAGGCTCTTTCAGTCCAGCAGTATGATCGGTATGTGTGTGAGTTCGACCTCGGAGAGTCT 2344
Qy 5677 TTGGGCTGTCTGGGTGTGAATGGAGCAGGGAAGAGCTCCAGTTTTCGATGTGTGACGGGG 5736
Db 2345 TTGGCTCTCTGGGAGTGAATGTGTGAGCAAAACCAACCATTCACAGATCTCACTGGGG 2404
Qy 5737 ACACATTTGGCAGCAGGGGGAGGCTGTGTGTGAGAGCCACAGGTCGCCCGGGAACCCA 5796
Db 2405 ACACCAAGTGAACATCGGGGATGCTACTGTGTGAGGCAAGAGCATTTAACTAGTATTT 2464
Qy 5797 GTGCTGGCACCTCAGATGGGATCTGCCCTCAATCCGATGCCATCTTTGAGCTGTGA 5856
Db 2465 CTGATGTCCATCAAAACATGGGCTACTGTCTCAGTTTGTATGCAATCGATACCTGTCA 2524
Qy 5857 CGGCGCCGAGCAGCCTCGAGCTGTCTGCGCGCTTGGCGGCTGTCCCGAGGCCCAGGTTG 5916
Db 2525 CAGCAGAGAAACCTTTTACTCTATGCCAGGCTGCGGGGCTGCCATCGAAGGAAATCG 2584
Qy 5917 CCAGACCGCTGGCTCAGGCTTGGCGCTGTGAGACTCTCATGGTACGACAGCCGCTG 5976
Db 2585 AGAAGTTGCAACCTGGGGTATCCAGAGCCTGGGCTTGTCTCTACGCTGACCGCTGG 2644
Qy 5977 CAGCACCTACAGCGGAGGAAACAAACGCAAGCTTGGCGCGCCCTGCGCTGTGGGG 6036
Db 2645 CAGGCACTACAGTGGAGGCAATAGAGGAACTCTCTACAGCCATAGTCTCTACTGGCT 2704
Qy 6037 ACCCAGCCGTGTGTCTTCTGAGCAGCCGACCAACAGGATGGACCCAGCGCGCGGCT 6096
Db 2705 GCGCTCCCTGCTGCTGTGATGAGCCACGACAGGATGGATCCCGAGSCACGCCGA 2764
Qy 6097 TCTTTTGAACAGCCTTTTGGCGTGTGCGGAGGCGGCTTCAGTGATGCTCACTCC 6156
Db 2765 TGTGTGGAACACCATTTGTGAGCATCATAGAGAGGAGAGCTGTGTCTCACTCC 2824
Qy 6157 ATAGCATGAGGAGTGTGAAGCGCTGTCTGCGGCTTAGCCATCATGTTGATGGGGGT 6216
Db 2825 ACAGCATGGAAGAAATGGAAGCTGTGTAGCGGCTGGCCATCATGTTGAGGGCACCT 2884
Qy 6217 TCCGCTGCTCTGGGAGCC--CGCAACATCTCAAGGGCAGATTTGCGGGGGTTCACACACTG 6275
Db 2885 TTCACTGTCTGGGTACCCATCCCAACACCTCAAGTACAAAGTTTGGAGAGCGCTACTATGTC 2944
Qy 6276 ACCGTGGGGTGGCCCGGCAAGTCCAG-----CCGGCAGCGGCC 6317
Db 2945 ACAATGAAATAAAATCTCCAAAGGACGACTGTGCTTCCGATTTGAAATCTCTGTGGAGCAG 3004
Qy 6318 TTCTGGCGGCGAGTTTCCCTGGGTGCGAGTGTGCGGAGGCACATGAGGCGCGCTGCGC 6377

Db 3005 TTCTTCCAGGCAATTCCTCGCAGCGTGCAGAGGAGAGACCCACAGCATGCTCCAG 3064
Qy 6378 TTCCAGCTCCCGCGGAGGGCGCTCGCCCTGCGCGGTCTTTGGAGAGCTGGCGGTG 6437
Db 3065 TTCCAGGTCCCTCG-----TCCTCCCTGGCCAGGATCTTCAGCTGCTCATTTCC 3115
Qy 6438 CACGGCGCAGAGACACGGCGTGGAGGACTTTTCCTGAGCCAGACGATGCTGGAGAGGTA 6497
Db 3116 CACAGGACAGCTGCTCATCAGGAGTACTCAGTCACCCAGACACAGCTGACACAGGTG 3175
Qy 6498 TTCTTGTACTTCCAGAGACACAGGGAGAGACGAGGAC 6536
Db 3176 TTTGTAATCTCGCTAAACAGCAGCAGTACAGACTATGAC 3214

RESULT 9
BQ955522 958 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT 8843758 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6456042
5', mRNA sequence.

ACCESSION BQ955522
VERSION BQ955522.1 GI:22371000
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 958)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: sgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: L1CM2635 row: k column: 19
High quality sequence stop: 517.
Location/Qualifiers

FEATURES
source
1..958
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6456042"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by King Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 11.7%; Score 799; DB 5; Length 958;
Best Local Similarity 93.2%; Pred. No. 1.4e-153;
Matches 892; Conservative 0; Mismatches 55; Indels 10; Gaps 5;

Qy 2982 CAGGATGTGGGGCTGCTCCAAAGCAGAGTGTGCAGACTCGCCACTCTCTGGTGGGATG 3041
Db 1 CAGGATGTGGGGCTGCTCCAAAGCAGAGTGTGCAGACTCGCCACTCTCTGGTGGGATG 60
Qy 3042 CAACGGAGCTGTCCGTGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTGTGAC 3101
Db 61 CAACGGAGCTGTCCGTGGCCATTGCCCTTTGTGGCGGCTCCCAAGTTGTTATCTGTGAC 120

Qy 3102 GAGCTACGGCTGGCGTGGATCCTCTTCCCGCGGGTATTTGGAGCTGCTGTCAA 3161
Db 121 GAGCTACGGCTGGCGTGGATCCTCTTCCCGCGGGTATTTGGAGCTGCTGTCAA 180
Qy 3162 TACCAGAGAGGTGCGACGCTGATCTCTCCACCCACCTCGATGAGGAGAGCTGCTG 3221
Db 181 TACCAGAGAGGTGCGACGCTGATCTCTCCACCCACCTCGATGAGGAGAGCTGCTG 240
Qy 3222 GGAGACCGTGTGGCTGTGTGGTGGAGTGGCGCTTGTGTGTGTGTGTGTGTGTGTGT 3281
Db 241 GGAGACCGTGTGGCGCTGTGTGGTGGAGTGGCGCTTGTGTGTGTGTGTGTGTGTGT 300
Qy 3282 CTGGCGCTGCTACCTGGGCTCCGGCTACTACCTGAGCTGGTGAAGGCCCGCTGCCCTG 3341
Db 301 CTGGCGCTGCTACCTGGGCTCCGGCTACTACCTGAGCTGGTGAAGGCCCGCTGCCCTG 360
Qy 3342 ACCACCAATGAGAGGCTGACACTGACATGAGGAGGAGTGTGGACACACAGGAGGAGAA 3401
Db 361 ACCACCAATGAGAGGCTGACACTGACATGAGGAGGAGTGTGGACACACAGGAGGAGAA 420
Qy 3402 AAGATGCGAGGAGGAGGAGAGTGGGACTCTCTAGCTGTGTGTGTGTGTGTGTGTGTGT 3461
Db 421 AAGATGCGAGGAGGAGGAGTGGGACTCTCTAGCTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 3462 TGGGTGCGCGGGGACGGCTGTGTGGAGAGTGGCCACAGAGCTGTGTGTGTGTGTGTGT 3521
Db 481 TGGGTGCGCGGGGACGGCTGTGTGGAGAGTGGCCACAGAGCTGTGTGTGTGTGTGTGT 540
Qy 3522 TACACGGGTGCCCATGACGGGAGCTTTCGCCACACTTTCGGAGAGCTAGACACGGGG 3581
Db 541 TACACGGGTGCCCATGACGGGAGCTTTCGCCACACTTTCGGAGAGCTAGACACGGGG 600
Qy 3582 GCGAGCTGAGGCTCACTGGGCTACGGGATCTCCGACACAGGCTCGAGAGATCTTCTCT 3641
Db 601 GCGAGCTGAGGCTCACTGGGCTACGGGATCTCCGACACAGGCTCGAGAGATCTTCTCT 660
Qy 3642 AAGTGTGTGGAGGAGTGTGTGGGACACAGATATGGAGGATGGGAGCTGGGGGAGCAG 3701
Db 661 AAGTGTGTGGAGGAGTGTGTGGGACACAGATATGGAGGATGGGAGCTGGGGGAGCAG 720
Qy 3702 CTATGCACAGCA-TTGTGTGGCTTAGAGCTAACCCCTGGGGCTCAAGATGCCGCCACAGA 3760
Db 721 CTATGCACAGGCAATTTGTGGCTTAAACCGTACCTTCCAGTCCAGATGCCGGCAGAGA 780
Qy 3761 GACAGCGCTGG--AGAACGGGGAACAGCTGGGTGAGCCCGCCAGAGACTGACAGGG--CT 3816
Db 781 GACAGCGCTGGGAGAACGGGGAACCAACTTGGTTCAGGCCCCAGAGACTGACAGGGGCTC 840
Qy 3817 CTGGGCCAGAGCGCGTGGGGCGGGGTACA---GGGCTGGGGCACTGACCCGCCAGAGCTCC 3873
Db 841 TGGGGCCAAACCCCGGGGGCGGGTTAACGGGGGTGGGGACCGAAACCCGCCACAGCTCC 900
Qy 3874 AGGCCCTGCTTCTCAAGCGCTTT--CTGCTTGGCGCGCAGCGCCCGGGCTGTT 3928
Db 901 AGGCCCTGCTTCTTAAAGGCTTTTCTGCTTTGCCCCCAACCCCGCGGGCGGCT 957

RESULT 10
BC036566
LOCUS BC036566 1934 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, clone IMAGE:1366813, mRNA.
ACCESSION BC036566
VERSION BC036566.1 GI:23306463
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1934)
Strausberg, R.
Direct Submission
Submitted (20-AUG-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbe@mail.nih.gov
 Tissue Procurement: Marcello Bento Soares, Ph.D.
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Lilaea Prabhu, Parvaneh Saesdi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 67 Row: b Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
 This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers
 1..1934
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1366813"
 /tissue_type="mammary gland"
 /clone_lib="Soares_mammary_gland_NbMNG"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac"

ORIGIN

Query Match 11.2%; Score 760.6; DB 3; Length 1934;
 Best Local Similarity 73.3%; Pred. No. 1.3e-145;
 Matches 1038; Conservative 0; Mismatches 304; Indels 75; Gaps 2;
 5348 TGAGCGCTGGGAGACAGGCGATTCCAGTCAACCCCTGCGTGGGAGGTGGTTCGCAAGAA 5407
 428 TGTCCCATAGGAGACAAAGCAATTTCAGTCACCCCTACGCTGGGACATCATTTGCAAGAA 487
 5408 CCTTTGGCGATGTGTGATACAGGGGCCCTCTTCCCTCTCTTCTACATCTGCTGCGAGCA 5467
 488 CCTCTGGCATGATGCGCCAGGGACCTCTGTCTCTCTCATCACTCTCTGCTCCCAAC 547
 5468 CCGAAGCCACTCTGCGACAGCCAGGGTG----- 5498
 548 CCGAACCGTCTCTGCGACAGTCAGTGGGAACAGGTGAGGCGAGGCGGCGCAGAGTC 607
 5499 -----AGTCTCTGCGCACTCCT 5515
 608 CTTCCAGGTCCTAACTGTCTATCTCTAGATCAAAACCAAGATGCTGCGCCCT 667
 5516 GGGAGAGAGACAGAGATGTAGCCGTGAACGGGAGCGGTGTGTCGAAGAGCAACCA 5575
 668 GGGGAGGAGATGAGATGTGGTCAAGAGCGTGAAGCGGTGACCAAGGGGGTACCCA 727
 5576 GGGGAGTGTGTGCTGAGGAATTTGACCAAGCTATACCGTGGGCGAGAGATGCCAGC 5635
 728 GGGGAGTGTGTAGTCTCAGGAGCTTGACCAAGTTTACCGTGGGCGAGAGAACCCAGC 787
 5636 TGTGACCGCTTGTGCTGGGATTTCCCTCTGCTGAGTGTGTTTGGGCTGCTGGGTGTGAA 5695
 788 TGTGATCGCTGTGCTTAGGATCCCCCTGGGGAGTGTTCGGGCTGCTGGGTGTCAA 847

RESULT 11
 BUS28031
 LOCUS

DEFINITION AGENCOURT_10247007 NIH_MGC_101 Homo sapiens cDNA clone
 BUS28031 914 bp mRNA linear EST 13-SEP-2002

5596 TGGACAGGGAAGACGCTCCACGTTTCCATGCTGTGACGGGGGACACATTTGGCCAGCAGGG 5755
 848 CGGGGACAGGGAAGACATCCACCTTCCCATGCTGTGACAGGGGACACATTTGGCCAGCAG 907
 5756 CGAGGCTGTGCTGGCAGGCGACAGCGTGGCCGGGAAACCCAGTGCTGCTCGCACCTCAGCAT 5815
 908 TGAAGCAGTACTGGCAGGTCAACAGTGGCCAGGAGCGTCTGCTCGGCGACCGCAGCAT 967
 5816 GGGATACTGCGCTCAATCCCATGCTCTTTGAGTCTGTAACGGGCGCGAGCACTGGA 5875
 968 GGGCTACTGTGCTCCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
 5876 GCTGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5935
 1028 ACTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087
 5936 CTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5995
 1088 CTTGGTGGCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1147
 5996 GAACAAACGCAAGCTGCGACGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6055
 1148 CAACAAACGCAAGCTGCGACGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207
 6056 GGACGAGCCACACAGGATGACCCAGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 6115
 1208 GGAAGAGCCACACAGGATGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267
 6116 GGCCTGCTGCGGAGGCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6175
 1268 GTGCTGCTGCGGAGGCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1327
 6176 AGCGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6235
 1328 AGCGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387
 6236 GCAACATCTCAAGGCGAGTTCGCGGCGGTGACACACTGCTGCTGCTGCTGCTGCTGCTG 6295
 1388 TCAGCATCTCAAGGCGAGTTCGCGGCGGTGACACACTGCTGCTGCTGCTGCTGCTGCTG 1447
 6296 AAGTCTCCAGCGCGCGCTTCTGCTGCGCGCGAGTTCCTGCTGCTGCTGCTGCTGCTGCTG 6355
 1448 CCAGCTGAGCGCGGATAGCTTCACTAGGACCCACATTCCTGCGGCTGAACTCCCGGA 1507
 6356 GGACATGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6415
 1508 GGTGACGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1567
 6416 CTTCTTTGGAGAGCTGCGGCTGCGCGGCGAGGACGCGGTGGAGGACTTTTCCTGCTGAG 6475
 1568 AGTGTTCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1627
 6476 CCAGACATGCTGAGGAGGTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6535
 1628 CCAGACCATCTGAGGAGGTGCTTCTATATTTCTCCAGAGCCACAGGGGAGAGGA 1687
 6536 CACCGAAGACAGAGGAGGAGGTGGAGTGGACCCCGCGCGAGCGCTGCGAGCACC 6595
 1688 GAGC---AGTCCGCGAGGAGCTGAAAGAGGAGGTTCCTCAAACTGCGCGCGCAGCATCC 1744
 6596 CAAACGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6655
 1745 CAAACGCTGAGCGGATTCCTGGAAGACCCAGCTCTGTGAGAGCCATGATCTGAGCATG 1804
 6656 CTTCCCTGCGGGCGCGCGGCGGCGCTGCGGAGTGG 6692
 1805 CTTCCCTGCGGCTGAGTGGCAAGCTCAGACAGAG 1841

IMAGE:6537923 5', mRNA sequence.

ACCESSION
BU528031
VERSION
KEYWORDS
SOURCE
ORGANISMBU528031.1 GI:22838472
EST.
Homo sapiens (human)REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2701 row: k column: 11
High quality sequence stop: 602.
Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6537923"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site:1: EcoRI; Site:2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 11.0%; Score 751.8; DB 5; Length 914;
Best Local Similarity 94.5%; Pred. No. 7.5e-144;
Matches 834; Conservative 0; Mismatches 42; Indels 7; Gaps 5;

QY 2982 CAGGATGGGGCTGCTCCAGCAGAGTGTGCAGACTCGCCACTCTCTGGTGGGATG 3041
Db 1 CAGGATGGGGCTGCTCCAGCAGAGTGTGCAGACTCGCCACTCTCTGGTGGGATG 60

QY 3042 CAACGGAAGCTGTCCGTCGTCATTCCTTTGGGGCGGCTCCCAAGTTGTTATCTCGAC 3101
Db 61 CAACGGAAGCTGTCCGTCGTCATTCCTTTGGGGCGGCTCCCAAGTTGTTATCTCGAC 120

QY 3102 GAGCTACCGCTGGCGTGGATCTCTGCTCCCGCCCGGTATTGGAGCTGTCTCTCAA 3161
Db 121 GAGCTACCGCTGGCGTGGATCTCTGCTCCCGCCCGGTATTGGAGCTGTCTCTCAA 180

QY 3162 TACCGAGAGCTCGCACCTGATCTCTCCACCCACCTGGATGAGGCGAGAGTGTCTG 3221
Db 181 TACCGAGAGCTCGCACCTGATCTCTCCACCCACCTGGATGAGGCGAGAGTGTCTG 240

QY 3222 GGAGACCTGTGGCTGTGGCGAGTGGCGCTTGTGCTGTGGCTCCCTCTTC 3281
Db 241 GGAGACCTGTGGCGAGTGGCGCTTGTGCTGTGGCTCCCTCTTC 300

QY 3282 CTGGCCCTGCTACCTGGCTCCGGCTTACTCTGAGCTGTGGAGGCGCGGCTGCCCTG 3341
Db 301 CTGGCCCTGCTACCTGGCTCCGGCTTACTCTGAGCTGTGGAGGCGCGGCTGCCCTG 360

QY 3342 ACCACCAATGAGAGCTGACACTGACATGATGAGGCGAGTGTGGACACACGAGGAGAAAG 3401
Db 361 ACCACCAATGAGAGCTGACACTGACATGATGAGGCGAGTGTGGACACACGAGGAGAAAG 420

QY 3402 AAGATGCGAGCCAGGGCAGCAGAGTCTGGCACTCTCAGCTGCTGGCCCTGTACAGCAC 3461
Db 421 AAGATGCGAGCCAGGGCAGCAGAGTCTGGCACTCTCAGCTGCTGGCCCTGTACAGCAC 480

QY 3462 TGGGTGCCCGGGGCGACGGCTGGTGGAGAGCTGCCACAGAGCTGGTGTGTGTGCTGCC 3521
Db 481 TGGGTGCCCGGGGCGACGGCTGGTGGAGAGCTGCCACAGAGCTGGTGTGTGTGCTGCC 540

QY 3522 TACACGGTGGCCATGACGGAGCTTCCGACACACTTCCGAGAGCTAGACACGGGCTG 3581
Db 541 TACACGGTGGCCATGACGGAGCTTCCGACACACTTCCGAGAGCTAGACACGGGCTG 600

QY 3582 GCGGAGCTGAGGCTCACCTGGCTACGGGATCTCCGACACAGCTCGAGGAGATCTT-CCT 3640
Db 601 GCGGAGCTGAGGCTCACCTGGCTACGGGATCTCCGACACAGCTCGAGGAGATCTTCCCT 660

QY 3641 GAAAGTGGTGGAGGAGTGTCTGGCGGACACAGATATGGAGGATGCGAGCTGCGGGCAGCA 3700
Db 661 GAAAGTGGTGGAGGAGTGTCTGGCGGACACAGATATGGAGGATGCGAGCTGCGGGCAGCA 720

QY 3701 CCTATGACACAGGATCTGCTGGCTAGAGCT-AACCTGGGCTCAAGATGCGGCAC-AG 3758
Db 721 CCTATGACACAGGATCTGCTGGCTAGAGCTAAACCTCAAGATGCGGCACCAAG 780

QY 3759 GAGACAGCG-CTGGAGAACGGGGAAC--AGCTGGGTGAGCCCGCAGAGACTGACACAGG 3814
Db 781 GAGACAGCGCTGGAGAACGGGGAACAGAGCTGGTGGTGGCGGCGGAGCTGACACAGG 840

QY 3815 CTCTGGGCGACAGCCGCTGGGCGGCTAGAGGCTGGGCGACTG 3857
Db 841 CTCTGGGCGCGAGACCCCGTGGGCGGCGGAGAGCGGCGCTTG 883

RESULT 12

BU846021

LOCUS

DEFINITION
AGENCOURT 10413059 NIH MGC 109 Homo sapiens cDNA clone
IMAGE:6579963 5', mRNA sequence.

ACCESSION

BU846021

VERSION

BU846021.1 GI:24030726

KEYWORDS

EST.

SOURCE

Homo sapiens

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2784 row: c column: 03

High quality sequence start: 23

High quality sequence stop: 473.

Location/Qualifiers

1..969

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6579963"

/tissue_type="teratocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 109"

/note="Organ: ovary; Vector: pOTB7; Site:1: EcoRI; Site:2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned

Db 554 CGCCTAGCATCATGCTGAATGGCGGTTCCGCTGCTGGGAGCCGCCAATCTCAAG 495
Qy 6249 GGCAGATTCCGCGCGGTCACACACTGA-CCCTGGGGTGGCCGCGAGTCCAGCC 6307
Db 494 GGCAGATTCCGCGCGGTCACACACTGA-CCCTGGGGTGGCCGCGAGTCCAGCC 435
Qy 6308 GGCAGCGGCTTCTGTCGGCGGCGAGTTCCTGGGTGGAGCTGCGGAGGACATGGAGG 6367
Db 434 GGCAGCGGCTTCTGTCGGCGGCGAGTTCCTGGGTGGAGCTGCGGAGGACATGGAGG 375
Qy 6368 CCGCTGCGCTTCCAGCTGCGCGGAGGCGCTGCGCCCTGGCGCGCTTTTGGAGA 6427
Db 374 CCGCTGCGCTTCCAGCTGCGCGGAGGCGCTGCGCCCTGGCGCGCTTTTGGAGA 315
Qy 6428 GCTGCGGCTGACGCGCGAGACACGCGCTGGAGGACTTTTCCTGAGCGACAGCATGCT 6487
Db 314 GCTGCGGCTGACGCGCGAGACACGCGCTGGAGGACTTTTCCTGAGCGACAGCATGCT 255
Qy 6488 GGAGGAGGTATTTCTTGTACTTCTTCCAGGACACGAGGAGGACGAGGACACCGAAGAGCA 6547
Db 254 GGAGGAGGTATTTCTTGTACTTCTTCCAGGACACGAGGAGGAGGACACCGAAGAGCA 195
Qy 6548 GAGGAGGAGGAGTGGAGTGGAGCCCGCGCGAGCTGCGAGCAACCCAAACGCTGAG 6607
Db 194 GAGGAGGAGGAGTGGAGTGGAGCCCGCGCGAGCTGCGAGCAACCCAAACGCTGAG 135
Qy 6608 CCAAGTTCTCGATGACCTTAGCACTGCGGAGAGTGTCTGAGCTCCCTCCCTGCGG 6667
Db 134 CCAAGTTCTCGATGACCTTAGCACTGCGGAGAGTGTCTGAGCTCCCTCCCTGCGG 75
Qy 6668 GCGCGCGGAGGCGCTTGGGAATGGCAAGGTAGAGTGCCTAGGAGCCCTGGACT 6727
Db 74 GCGCGCGGAGGCGCTTGGGAATGGCAAGGTAGAGTGCCTAGGAGCCCTGGACT 17
Qy 6728 CAGGCTGCGAGAGGG 6743
Db 16 CAGGCTGCGAGAGGG 1

RESULT 14
CD629699/c 749 bp mRNA linear EST 12-JAN-2004
LOCUS 55052443J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD629699
ACCESSION CD629699
VERSION CD629699.1 GI:40277965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Fu.G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1..749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 10.5%; Score 717.4; DB 6; Length 749;
Best Local Similarity 98.9%; Pred. No. 8.7e-137;
Matches 733; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

Qy 4969 ACATGTGTAACTACTTGGTGCCA--GCATGCATGTGGTGTCTCATCTTTCTGSCCTTCCA 5026
Db 741 ACATGTGTAACTACTTGGTGCCAAGCATCTCTGGTGTCTCATCTTTCTGSCCTTCCA 682
Qy 5027 GCAGAGGCGATATGTGGCCCTTGCACACCTGCTGCTCTCTCTGCTGTCTACTACTGTA 5086
Db 681 GCAGAGGCGATATGTGGCCCTTGCACACCTGCTGCTCTCTCTGCTGTCTACTACTGTA 622
Qy 5087 TGGCTGTGCGATACACCGCTCATGTACCCAGCGCTCTCTTCTCTCGTCCGCCAGCAC 5146
Db 621 TGGCTGTGCGATACACCGCTCATGTACCCAGCGCTCTCTTCTCTCGTCCGCCAGCAC 562
Qy 5147 AGCCTATGTGGTGTCTCACCTGCATAAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCAC 5206
Db 561 AGCCTATGTGGTGTCTCACCTGCATAAAACCTCTTTATTTGGCATCAATGGAAGCATGGCCAC 502
Qy 5207 CTTTGTGTGTGAGCTCTTCTGTATCAAGAGCTGACGAGGTGAGCCGATCTTGAACA 5266
Db 501 CTTTGTGTGTGAGCTCTTCTGTATCAAGAGCTGACGAGGTGAGCCGATCTTGAACA 442
Qy 5267 GGTCTCTCTTATCTTCCCGCACTTCTGCTTGGGCGGGGGCTCATGACATGGTGGGAA 5326
Db 441 GGTCTCTCTTATCTTCCCGCACTTCTGCTTGGGCGGGGGCTCATGACATGGTGGGAA 382
Qy 5327 CCAGGCCATGGTGTATGCTTGTAGCGCTTGGGAGACAGGCAAGTTCAGTCAACCCCTGCG 5386
Db 381 CCAGGCCATGGTGTATGCTTGTAGCGCTTGGGAGACAGGCAAGTTCAGTCAACCCCTGCG 322
Qy 5387 CTTGGAGTGTGCGCAAGAACCTCTTGGCCATGTGTATACAGGGGCGCCCTCTTCTCT 5446
Db 321 CTTGGAGTGTGCGCAAGAACCTCTTGGCCATGTGTATACAGGGGCGCCCTCTTCTCT 262
Qy 5447 CTTACACTACTGCTGCGAGCACCGAGCAACTCTGCGCACAGCCAGGCTGAGTCTCT 5506
Db 261 CTTACACTACTGCTGCGAGCACCGAGCAACTCTGCGCACAGCCAGGCTGAGTCTCT 202
Qy 5507 GCCACTCTCGGAGAGGAGGAGGATGTAGCCCTGAAACCGGAGCGGTGCTCCAAGG 5566
Db 201 GCCACTCTCGGAGAGGAGGAGGATGTAGCCCTGAAACCGGAGCGGTGCTCCAAGG 142
Qy 5567 AGCCACCAAGGGGATGTGTGGTGTCTGAGGAACCTTGACCAAGGTATACCGTGGGAGAG 5626
Db 141 AGCCACCAAGGGGATGTGTGGTGTCTGAGGAACCTTGACCAAGGTATACCGTGGGAGAG 82
Qy 5627 GATGCCAGCTTGTACCGCTTGTGCTGGGAGTTCCTGCTGCTGCTGCTGCTGCTGCT 5686
Db 81 GATGCCAGCTTGTACCGCTTGTGCTGGGAGTTCCTGCTGCTGCTGCTGCTGCTGCT 22
Qy 5687 GGGTGTGAATGGAGCAGGAA 5707
Db 21 GGGTGTGAATGGAGCAGGAA 1

RESULT 15
CD629709
LOCUS 5505387J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD629709
ACCESSION CD629709
VERSION CD629709.1 GI:40277975
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Fu.G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102

Email: gfu@incyte.com.
Location/Qualifiers
1. .730
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

FEATURES
Source

ORIGIN

Query Match 10.4%; Score 704.8; DB 6; Length 730;
Best Local Similarity 99.5%; Pred. No. 3.4e-134;
Matches 728; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 3395 GGAAAGAAGAAATGGCAGCCAGGCGACAGAGTCGGCACTCTCAGCTGTGGCCCTGGT 3454
Db 1 GGAAAGAAGAAATGGCAGCCA-GGCAGCAGAGTCGGCACTCTCAGCTGTGGCCCTGGT 59

Qy 3455 ACAGCACTGGGTGCGCGGGCAGCGCTGGTGGAGAGTGCACACGAGCTGGTGGT 3514
Db 60 ACAGCACTGGGTGCGCGGGCAGCGCTGGTGGAGAGTGCACACGAGCTGGTGGT 119

Qy 3515 GCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTAGACAC 3574
Db 120 GCTGCCCTACACGGGTGCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTAGACAC 179

Qy 3575 GCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCACACCGCTTCGAGGAGAT 3634
Db 180 GCGGCTGGCGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCGCTTCGAGGAGAT 239

Qy 3635 CTTCTGAAGTGTGGAGGAGTGTCTGGCGACACAGATATGGAGATGGCAGCTGGG 3694
Db 240 CTTCTGAAGTGTGGAGGAGTGTCTGGCGACACAGATATGGAGATGGCAGCTGGG 299

Qy 3695 GCAGCACTATGCACAGCATTTGTGGCTAGACGTAACTTGGCGCTCAAGATGCCGCC 3754
Db 300 GCAGCACTATGCACAGCATTTGTGGCTAGACGTAACTTGGCGCTCAAGATGCCGCC 359

Qy 3755 ACAGGAGACCGCTGGAGAAACGGGAAACCAAGCTGGGTACGCCCGCAGAGCTGACCGGG 3814
Db 360 ACAGGAGACCGCTGGAGAAACGGGAAACCAAGCTGGGTACGCCCGCAGAGCTGACCCA-GG 418

Qy 3815 CTCTGGGCGCAGCGCGCTGGCGGGTACAGGGCTGGGCACTGACCGCGCAGCAGCTCCA 3874
Db 419 CTCTGGGCGCAGCGCGCTGGCGGGTACAGGGCTGGGCACTGACCGCGCAGCAGCTCCA 478

Qy 3875 GGCCCTGCTTCTCAAGCGCTTTCTGTTCGCCCGCAGCGCGCGCGCTGTTCGCCCA 3934
Db 479 GGCCCTGCTTCTCAAGCGCTTTCTGTTCGCCCGCAGCGCGCGCGCTGTTCGCCCA 538

Qy 3935 GATCGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCTGTCCTCC 3994
Db 539 GATCGTGCTGCTGCCCTCTTTGTGGGCTGGCCCTCGTGTTCAGCCTCATCTGTCCTCC 598

Qy 3995 TTTTCGGGCACTACCGGCTCTGGCGCTCAGTCCACCATGTACGGTCTCAGGTGTCTT 4054
Db 599 TTTTCGGGCACTACCGGCTCTGGCGCTCAGTCCACCATGTACGGTCTCAGGTGTCTT 658

Qy 4055 CTTTCAGTGAGGACGCGCGCGGACCTTGAGACGTGCCCGCTGTTCGAGGCGCTGCTGCA 4114
Db 659 CTTTCAGTGAGGACGCGCGCGGACCTTGAGACGTGCCCGCTGTTCGAGGCGCTGCTGCA 718

Qy 4115 GGAGGCAGGACT 4126
Db 719 GGAGGCAGGACT 730

Search completed: December 30, 2004, 05:09:45
Job time : 18920 secs